

Conceptual sovereignty problems

The British and Soviet governments are beset by problems of sovereignty, as it happens the obverse of each other. They (and other governments) should follow the dictum that national interests are but an amalgam of what people want.

WHEN absolute monarchs were commonplace, sovereignty was a simple matter. The monarch would say when a threat of external interference was intolerable, and would conscript an army of serfs to resist it. The rise of representative democracy has muddled this simple concept. That much is clear from events this week, notably Mr Mikhail Gorbachev's continuing search for a union treaty that will reconcile the interests of the republics that constitute the Soviet Union with those of what is left of the central government. Then, at the weekend, Mr John Major, Britain's new prime minister, will be seeking a conciliatory way of squaring British national interests with the European Communities' plans for a European central bank, provisionally called Eurofed. No doubt his main argument will be that the proposal puts the economic cart before the horse, but there will also be references to British sovereignty. Of what does that now consist?

Confusion between sovereignty and national identity is a persistent nuisance now that the citizens of nation-states commonly differ from each other in language, religion and even (but less commonly than supposed) in ethnic origin. The sovereign state of Switzerland accommodates four languages (counting Romanche) in a population of merely seven million, yet the Swiss are generally agreed that the sovereign territory will be vigorously defended against increasingly unlikely military attack. Britain may go the same way if Ulster is pacified and if plans materialize to devolve regional government to Scotland and, eventually, to Wales. But that would be no more inconsistent with British sovereignty than would an acceptable division of authority between central and regional governments in the Soviet case.

At the outset of the British liberal tradition, David Hume remarked that national interests are but the sum of individuals' interests, but he was before his time. He reckoned without the ways in which national governments seek to blur the distinction between sovereignty, which is the legitimate expression of national interest, and the general sense of national identity. Hume could not have guessed how modern governments would cultivate a spurious sense of national identity by devices such as subsidized national airlines, television networks under orders to foster cultural identity and even support for international sports competitions. More old-fashioned and reprehensible devices also persist. The classic stratagem of reinforcing the defence of national sovereignty

by emphasizing, exaggerating or even inventing an external threat (as now practised by Iraq) sustained US defence budgets and held the Soviet Union together through the Cold War.

Yet in the end, Hume will be proved right. That is how Britain and the rest of the European Communities should approach the weekend's negotiations. All 12 members are committed to a true single market from the beginning of 1993. They have also (except for Greece and Portugal) agreed that their currencies should retain the same relative value from now on, but may be devalued or revalued in exceptional circumstances. So why not also now settle for the goal of a common currency, administered by a central bank? On the face of things it should not matter.

But it does. British diffidence on the issue has not been sufficiently understood. The European Communities' internal arrangements do not yet allow for the redistribution of wealth between member-states except through the arbitrary Common Agricultural Policy and the much smaller Social Fund. So a member state suffering a long run of economic disappointment might well become socially impoverished as well. Increased personal taxation would not be a remedy, but would merely (in a true free market) drive the taxpayers elsewhere. In the event, an ill-prepared shotgun federation might be the only workable solution. If that is the cause of the British government's reluctance over the proposed single currency, it is legitimate in Hume's sense, and should be declared as such. By contrast, supposed British fondness for the pound sterling is no more substantial than national pride in the national airline, long since privatized.

Gorbachev's difficulty is the obverse of the British. He hopes to keep an existing federation intact. Again Hume's definition applies. Before Gorbachev's arrival on the scene, when there seemed no alternative to continued federation of the Soviet Union, and when even the thought that there might be one was rigorously suppressed, the republics toed the central line. Now they are busily calculating where their national interests lie.

What, if any, are the benefits of continued federation? Gorbachev can offer common defence against the outside world, but at the price of continued conscription into the armed services; the experience of the past few months suggests that inter-republic peace will not be as readily delivered. And such economic benefits as federation might bring must still seem elusive to the republics that



have suffered from Moscow's mismanagement in past decades, yet they are stuck with the distribution of capital resources brought about by 70 years of central planning, much of it skewed against independence. Yet Gorbachev's immediate task would be simplified, and the long-term interests of the republics would be more readily calculated, if national interests were not so often clouded by misplaced assumptions about national identity. □

Set-asides in health?

The British government, embarrassed by the cost of health care, should devise a new strategy for research.

RESEARCH never makes medical treatment cheaper, only ever more expensive, which no doubt partly explains the British government's antipathy towards it. And anybody who thinks that politics is not mostly about money — who it is taken from, and to whom it is given — has not been paying sufficient attention to recent history. The lesson the British government has been trying to teach its taxpayers for the past decade is that everything is a commodity. But the medical research community seems to have been particularly unwilling to learn. It has insisted not only on producing more of its own particular commodity than the government wishes to pay for through the National Health Service, but it also has had the nerve to seek to persuade the government to help produce even more by increasing support for research.

The problem is exacerbated because the health service itself spends virtually nothing on research, perhaps less than 2 per cent of the £1,400 million a year spent on health care research in the United Kingdom. The rest is spent by a motley collection of government agencies, charities and industry — which nevertheless have one thing in common: in general, they have not traditionally seen it as their problem to shoulder the costs of implementing the results of the research they sponsor. It seems to have come as a shock to many of them now to discover that the government will not let this attitude continue in perpetuity.

The situation is comparable with that in British agriculture. Encouraged by government subsidies, themselves born of the belief during the Second World War that Britain should produce as much food as possible, farmers now produce too much. So, earlier this year, the British government followed new EC regulations and, in an attempt to cut food production, introduced various set-aside schemes. The hope is to achieve substantial cuts in beef, lamb and crop production simply by paying farmers to produce less. It is an entirely voluntary scheme being tried out in different parts of Britain.

Will the government now be tempted to follow the same principles in medical research? Prevention being better than cure is both a traditional and a contemporary bromide. Its appeal for the present government is that, for many diseases, it is probably also cheaper. So why not pay research organizations to shift the emphasis of their

research from treatment to prevention? Nowhere would this make more sense than with cancer, where epidemiologists widely believe that as many as 80 per cent of all cancers are environmentally triggered, yet little is known about the specific triggers. And virtually nothing is being done to develop preventive strategies beyond vague exhortations to live sensibly.

Interventionist strategies might be even better. Here is an example. The prophylactic use of the oestrogen antagonist tamoxifen to prevent breast cancer may well be the equivalent for cancer of fluoride in the prevention of dental caries. But the point needs proving by clinical trials, and for such research to be effective, long-term commitments of funds would be required extending far beyond what the average charity or pharmaceutical company would be prepared to stomach. Yet an agreement to underwrite such a venture in exchange for a major switch in research emphasis by cancer charities in the early stages of the trial might make good economic as well as medical sense. □

Research at law

There is no immediate prospect that researchers will be sued for what they publish, but care is called for.

At least for the time being, most researchers should not take fright that apple-growers in the United States, who have been using 'Alar' to protect their crops against fungal infection, plan to sue the Natural Resources Defense Council (NRDC) over an allegedly damaging report published last year (see *Nature* 348, 383; 29 November 1990). On the face of things, there may seem a danger that publications critical of established techniques may generally expose their authors to damaging damage suits. But the circumstances of the NRDC report are not those attending most scientific publications. For one thing, the report was published by the organization itself, not by a journal with pretensions to impartiality. For another, although NRDC claims that its report was subjected to peer-review, only a trial will tell whether the 12 assessors were asked questions likely to elicit critical responses.

Yet the threat is not entirely negligible. While the bulk of the scientific literature is innocuous, there are occasions when a scientific article may seem to readers to be damaging to established interests, perhaps those of medical practitioners or manufacturers. Plainly an improvement of technique, whatever the commercial consequences, could not be held to be damaging. Even a demonstration that an existing technique has unexpected defects or consequences would also be so held. But occasions arise when the criticism implied by a piece of research is nearer the bone. Then the defence of publication must rest on the good faith of the authors and the conduct of the refereeing process. Luckily, even where the first amendment of the US constitution does not apply, nobody need fear that damage suits are indefensible. □

NIH propose cost benefit rules in grant reforms

- Research per dollar yardstick
- Total new grants limited to 5,750 per year

Washington

In the face of pressure from Congress and their own grant recipients to stretch research dollars, the US National Institutes of Health (NIH) have unveiled plans to change the rules dramatically in federally funded biomedical research.

On their way out are some archaic funding traditions — such as the quaint practice of ‘approving’ almost all applications regardless of whether they are actually funded or not — and some more modern rituals, such as the increasingly common tendency to chop a set percentage off all grants when money gets tight. In their place, NIH will introduce more case-by-case analysis of applications, based on their costs and some measure of the research per dollar they are likely to produce, and some strict limits on grant growth.

Draft copies of the plan were distributed to science policy-makers last week, and a two-day hearing on the reorganization will be held at NIH starting on 17 December.

According to the plan, NIH propose to:

- Begin considering the costs of each proposal, rather than its scientific merit alone. Specifically, ‘indirect costs’ — the overhead percentage each university takes off the top of its researchers’ grants — will for the first time be considered in selecting grants for approval.

- Restrict the growth of a grant so that it increases no faster than the actual cost of research, as determined by the Biomedical Research and Development Price Index. Between 1980 and 1989, the average cost of an NIH project grant increased 16 per cent more than the index, according to NIH statistics.

- Adjust individual awards to maintain the total number of new grants at 5,750 a year. With a grant length of four years, that should allow for a total pool of 23,000 NIH grant recipients.

- Abandon the practice of ‘approving’ 95 per cent of applications, while less than a quarter of those are actually funded. Instead, NIH plan to adopt a system where all potentially fundable applications receive a score as well as a percentile ranking. Generally, there would be a cut-off level based on the current availability of funds, and those grants with rankings above that level would be funded. But because NIH intend increasingly to consider the costs of proposals, more grants will be funded ‘out of order’ of

their research quality ranking.

- Make four years the average grant length. In recent years, it has lengthened to 4.3 years, with no end in sight, as researchers complained that shorter periods forced them to spend too much time writing grant applications and not enough time doing research. But longer grant periods have also meant that fewer new grants have been available in recent years, a situation that panicked the research community.

- Stop ‘downward negotiations’ — the process of imposing across-the-board reductions in grants when NIH run out of money. Future cuts will be on a case-by-case basis, with the specific methods decided by each institute.

Many of the changes are intended chiefly to placate Congress. Pointing out that the cost of biomedical research has been rising faster than virtually all other economic indicators, Congress this year told NIH to regain control of the grant process by imposing a strict set of new financial constraints. The 1991 appropriations bill required NIH to report back with a plan this month.

Among the most controversial aspects of the new plan are the emphasis on indirect costs, and the case-by-case cuts when NIH find themselves overextended.

While ‘downward negotiations’ were the bane of many NIH-funded researchers, they were at least equally applied. Although NIH have not decided exactly how to replace the process, they are committed to discriminating between those that can safely take a large cut and those that cannot. Ideally, that will spread the cuts around as painlessly as possible. But it may also spread a good deal of confusion. “The scientific community may begin to look at downward negotiation as the good old days. At least they understood how it worked”, says John Diggs, NIH deputy director for extramural research, who put the plan together.

University officials are likely to be especially concerned by the threatened reduction in the amount they can recover from their researchers’ grants. Indirect costs, which pay for university construction, maintenance and incidentals, have increased by nearly 25 per cent over inflation in the past decade, nearly three times the rise in the cost of the research itself. The implication that universities are getting fat on research grants prompted Congress specifically to target indirect costs for scrutiny. Although NIH have not

decided exactly how to take relative indirect cost rates into consideration when comparing applications, some discrimination appears inevitable. Because they tend to have higher cost rates (up to 80 per cent in some cases), “private institutions are going to have a real problem”, notes Diggs. Universities on the coasts, where prices are high, and in climates that require considerable heating or cooling, are also likely to find themselves at a disadvantage, he says.

Another worry is the predetermined size of the research pool. Although Congress set the target of 5,750 new grants a year to placate researchers who feared that real numbers might dip far lower, NIH officials are concerned that such a target may prove restrictive. “The fear is that we will reach a plateau beyond which it will be difficult to grow”, says Diggs. In the past decade, the total number of awards has increased just 12 per cent, while the number of applications has more than doubled.

Christopher Anderson

Biologists lash back

Washington

BIOLOGISTS clashed with the Washington science establishment last week, rejecting the findings of a two-year study carried out by the National Academy of Science’s Institute of Medicine (IOM).

The IOM study suggest that, in the absence of real future growth in the budget of the National Institutes of Health (NIH), some money should be shifted from research to training grants (see *Nature* 347, 413; 4 October 1990). Thomas Edgington, president of the Federation of American Societies for Experimental Biology (FASEB), claimed this would cost the research community some 2,000 grants over the next ten years.

Training more biologists will only make the funding problem worse Edgington said. He cited a Princeton University report* that projects a surplus of some 1,700 biology PhD recipients a year by 2000. He also quoted a study from the *Chronicle of Higher Education* (28 March 1990), showing that while university faculty had increased by 6 per cent between 1975 and 1985, academic support personnel had increased by over 61 per cent in the same period — a decade in which the university overhead percentage of research grants also rocketed, he noted.

The IOM promptly released a statement defending its report and accusing FASEB of manipulating the numbers. Only 1,000 grants — out of the total in a decade of nearly 60,000 — would be affected, IOM said.

Christopher Anderson

* *Prospects for Faculty in the Arts and Sciences* (Princeton University Press, 1989).

Rows floor factory project

Tokyo & Washington

WHEN Hiroyuki Yoshikawa of the University of Tokyo proposed the Intelligent Manufacturing System (IMS) project to develop automated factories of the future, his dream was to redress the imbalances in world trade by systematizing manufacturing technology so that it can be transferred easily to developing nations. But instead, IMS is turning into a struggle between the economic giants of the world over control and organization of the yet-to-be realized project.

Representatives of the US Department of Commerce (DOC), Japan's Ministry of International Trade and Industry (MITI) and the European Communities (EC) at a recent meeting in Tokyo agreed on a strategy for launching a feasibility study for IMS early next year. However, fundamental differences remain between the three sides on how to organize the project and in particular how to handle intellectual property rights arising out of the research.

Under IMS, international research consortia drawn from government, industry and academic institutions will develop standardized manufacturing systems that are fully computerized from the design of a product through to its retail and distribution. In addition to dozens of Japanese companies that have signed up for the project, many of the world's largest multinational companies have also expressed interest in joining, including Rolls-Royce, Fiat, Philips and Siemens from Europe, and Rockwell International, United Technologies, Pratt and Whitney, IBM, DEC, Kodak, AT & T, General Motors and General Electric from the United States. And although it is government officials who are discussing the organization of the project, much of the input comes

from industry representatives.

Top of the agenda at last month's meeting in Tokyo was a US proposal on how IMS should be run, which follows earlier proposals from Japan and the EC (see *Nature* 347, 320; 1990). A detailed outline of the US proposal, still confidential, shows that on the contentious issue of intellectual property rights, the United States puts strong emphasis on the region of origin of any invention made under IMS so that US inventions will benefit US companies first and foremost.

Japan, on the other hand, in a proposal issued in February, calls for a significant part of intellectual property rights to be held by a central IMS organization to which all IMS participants from all countries would have access in much the same way as Japanese companies share access to inventions resulting from national MITI projects. The EC in turn have put forward detailed rules that would assign rights according to the relative contributions of companies to an invention regardless of where it is made.

Another contentious issue is the areas of research on which the project will focus. MITI officials say they want to cover all areas of automated manufacturing. But the US government and the EC suspect that Japan is targeting US and European expertise in software, artificial intelligence and systems integration.

The draft US proposal circulated in Washington a few months ago calls for the United States to develop a 'lightning rod' IMS proposal that focuses attention on areas in which US industry is weak but Japan and Europe are strong, for example process technology and precision tooling and engineering.

And, according to participants at the Tokyo meeting, the US representatives were particularly insistent about US interest in these areas. In contrast, the EC position is that only partners of "equal strength" should participate in IMS projects.

Also yet to be resolved is the mechanism of funding. Japan proposed a central IMS organization funded with contributions from all participants, and such an organization is already in place in Tokyo with a fund of several million dollars donated from dozens of Japanese companies. The US proposal also calls for a central fund. But the EC insist that IMS should be a decentralized project with each region and each consortium funding its own research, much as in the case of the EC's ESPRIT project. All of these issues will have to be tackled in next year's feasibility study.

Hidehiko Nishiyama, deputy director of MITI's industrial machinery division, hopes to start some pilot IMS projects in

approximately eight months.

Further complications arise with other countries that wish to join. Canada, Australia and the European Free Trade Association were represented by observers at the Tokyo meeting and are expected to participate in the feasibility study as independent members. And the United States is pushing for IMS to be open to even more countries.

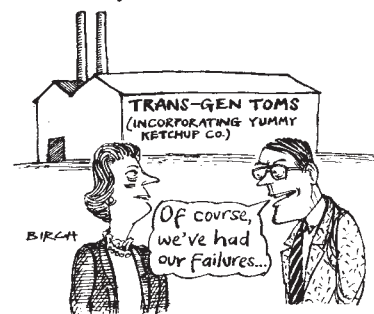
David Swinbanks & Chris Anderson

AGRICULTURAL BIOTECHNOLOGY —

Tables laid for transgenic tomatoes

San Francisco

DEVELOPERS of genetically engineered plants brought their products one tentative but important step closer to the American dinner table last week when Calgene, Inc. asked the US Food and Drug Administration (FDA) for a preliminary opinion on the use of a selectable genetic marker. This is the first time the FDA has been asked to evaluate a component of engineered plants which are to be consumed directly as food.



Although the gene, which confers plant cell resistance to the antibiotic kanamycin, is only a tool to help identify the sites of other genes introduced, its approval is expected to speed up future approval of engineered food crops. The gene was recently approved for human gene therapy and is being used in clinical tests.

Calgene, of Davis, California, wants to use the gene in tomatoes altered for enhanced flavour, and in genetically engineered cotton and rapeseed. Calgene's chief executive officer, Roger Salquist, claims this is the first such proposal submitted and will initiate the FDA's study of this new domain.

Calgene intends to file for approval of other products in the next 6-12 months for review in parallel with the kanamycin-resistance gene proposal, and it expects a decision within two years. In the wake of Calgene's move, other biotechnology companies are likely to submit plant products for approval. The company set a precedent before when it was the first to be granted permission for field trials by the US Department of Agriculture.

Elizabeth Schaefer

ISRAEL-SOVIET UNION

Working together

Jerusalem

ISRAEL and the Soviet Union have signed a scientific agreement in Jerusalem that paves the way for three years of joint research.

Last week's signing ceremony capped a week-long visit by a 16-member delegation from the Soviet Academy of Sciences, headed by its president, Professor Gury Marchuk.

Israeli Science and Technology Minister Yuval Ne'eman, himself a renowned physicist, called the agreement a "real breakthrough" in the scientific field and in the nature of diplomatic relations between the two countries. The Soviet Union broke off relations with Israel in 1967.

Lisa Perlman

SOUTH AFRICA

ANC moves on science policy

Johannesburg

THE African National Congress (ANC) held its first conference on science and technology here two weeks ago, at the University of the Witwatersrand. Described by Mohammed Valli Moosa, head of the ANC's political committee, as "successful far beyond expectations", the conference was meant to discuss the development of a national science and technology policy with as wide a cross-section of opinion as possible.

A year ago, few people in South Africa would have foreseen the presidents of the Council for Scientific and Industrial Research (Brian Clark), the Foundation for Research Development (Rein Arndt) and Atomic Energy Corporation chief Waldo Stumpf, discussing science and technology policy with Valli Moosa and Max Sisulu (head of the ANC's economic policy department). Other delegates included representatives from the Electricity Supply Commission (ESCOM), the universities and the private sector.

But even more remarkable was the degree of consensus. Valli Moosa outlined the ANC's draft position paper on science and technology policy (prepared by its interim science and technology group) in his opening address, but emphasized that it was mainly intended as a framework for discussion.

One of Moosa's prominent themes was the need to transform technology policy from its past emphasis on defence and the oil-from-coal industry to applications required for an improvement of living standards of South Africans or directed towards goods that could be produced competitively for international markets. The ANC, he said, is "acutely mindful that economic growth is to a great measure driven by advances in technology."

The ANC regards electrification as a technological priority, because of health problems arising from wood-smoke pollution in the townships, the expectation that electric light would enable people to study at night and the ecologically disastrous consequences of wood-collection. Yet the problems associated with electrification of the townships are not technological, but financial. This task would require an estimated R6,000 million, but no increased electricity generation, as ESCOM's existing power stations have a surplus capacity of 30 per cent.

A more intractable problem is the SO₂ emission in the Eastern Transvaal Highveld, where 80 per cent of South Africa's electricity is produced. Deposition amounts to 30–40 tonnes per square kilometre, matching the worst conditions found in Europe, but the cost of fitting scrubbers to each of thirty-two coal-fired power stations would increase the cost of

electricity by up to 50 per cent, making it even less affordable to the poor.

There was also concern at the conference about the declining proportion of South Africa's gross domestic product (GDP) spent on research and development (see *Nature* 344, 185; 1990). The conference urged a commitment to spending a certain percentage of GDP on research and development. David Jacobson of Altron, the South Africa's largest electronics corporation, made an appeal for tax incentives for research and development spending by the private sector. Theoretical physicist Rob Adam, of the ANC interim science and technology group, assured delegates that the ANC did not advocate "killing" fundamental research, as it recognized its importance.

The shortage of scientific and technical manpower in South Africa, and its racial composition, also occupied the conference. Ninety-six per cent of the engineers and 89 per cent of scientists in South Africa are white, largely as a consequence of an education system that produces one mathematics and science matriculation exemption for every 10,000 black school entrants.

While there are (badly paid) jobs aplenty for science graduates in primary and secondary schools, some debate centred around the extent to which a shortage of technical expertise is currently limiting industrial growth. One black South African who had recently returned to the country with a PhD in biochemistry from the University of California at Santa Cruz complained that he was unable to find a job in the private sector.

Professor C. C. Mjojo, president of the Network of African Scientific Organisations, in a closing address, explained how the network is attempting to convince African heads of state of the importance of science and technology in the development of the continent. It has already persuaded President Daniel Arap Moi of Kenya to convene a summit of 16 leaders to focus attention on the issue. Mjojo added that his organization was awaiting a signal from the ANC to begin talking to South African scientists. Valli Moosa responded that the ANC's isolationist policy had been directed at the "apartheid regime", and that Mjojo's presence showed that interaction had already begun.

Meanwhile, last week, Nelson Mandela received honorary degrees from both the University of the Western Cape and the University of Cape Town. Ironically, the graduation ceremony at UCT, held on the university's rugby field to accommodate all the guests, was presided over by the statue of Cecil John Rhodes gazing towards the hinterland. **Michael Cherry**

ANIMAL RIGHTS

Monkey case in Supreme Court

Washington

IN a surprising twist to the long-running and tortuous court battle over the 'Silver Spring monkeys'—the *cause célèbre* of the US animal-rights movement—the US Supreme Court agreed last week to hear arguments on a key procedural issue in the case, a move that could open yet another bloody chapter in the history of the research primates.

Since the monkeys first surfaced in 1981 as the focus of an undercover investigation by Alex Pacheco, who later went on to co-found People for the Ethical Treatment of Animals (PETA), animal-rights activists have maintained an unrelenting legal war on their behalf.

But until the Supreme Court decided to hear the case, it seemed that the activists were nearing the end of their legal tether. A suit, filed in Louisiana state court, charging the National Institutes of Health (NIH) with improper handling of the case, was referred last year to federal court, where it was dismissed. The activists asked the Supreme Court to rule on the propriety of that transfer, but had little hope that the court would take the case. But last week's decision raises the possibility of yet another round of legal battling, perhaps the fiercest yet. If the court sides with PETA, the case will go back to state court, where, for the first time, a judge will review the claims against the NIH, rather than simply procedural issues.

Christopher Anderson

ASTRONOMY

No end of blame

Washington

TECHNICIANS working on the Hubble Space Telescope (HST) knew that something was wrong with the shaping of the main mirror, but did not tell their supervisors, a NASA investigative team revealed last week. On at least three occasions, tests showed clear signs of a spherical aberration in the mirror, but the results were blamed on faulty test instruments and senior management was kept in the dark, said Lew Allen, former NASA administrator and chairman of the investigative panel.

The main error arose when a small fleck of anti-reflective paint chipped off a highly sensitive optical metering device, leading to a false reflection during tests. Other instruments recorded the error, but because they were less sensitive, it was discounted, Allen said. The technicians knew that the mirror was not being built as designed, but "the fact that there were errors did not reach anyone outside of the fabrication division". The cost of correcting the fault is estimated to be \$40–50 million.

Christopher Anderson

Western help comes slowly

Laxenburg, Austria

AN expert task force on the environment in Eastern Europe is urging governments and potential Western benefactors to concentrate first on so-called 'hot spots' that threaten public health before getting on with the monumental clean-up of polluted air, water and land in the region.

East European governments should focus their first efforts on building credible institutions for environmental protection equipped to enforce their policies before dealing with the problems.

The task force, which met last month at the International Institute for Applied Systems Analysis (IIASA) in Vienna, was the first meeting of scientific experts from all countries in the region. Its recommendations will be sent to East European environment ministers and potential donors and lenders such as the European Communities and the World Bank.

The overall picture of the environment that emerged at the meeting was unre-

New international organization

Paris

PLANS to establish an international environment group supported by 25 nations were approved at the World Bank headquarters in Paris last week.

The Global Environment Facility (GEF) which will have an estimated budget of more than \$1,000 million, will aim to promote "prudent environmental management", by encouraging 'green' technologies and strengthening policies put forward by individual countries.

The inspiration for GEF came from the French Finance Minister, Pierre Bérégovoy, with support from the German Federal Republic at the World Bank's annual meeting in September 1989. The Group should be running by the middle of next year.

The new organization will be implemented jointly by the World Bank, the United Nations Environment Programme (UNEP) and the United Nations Development Programme (UNDP).

It will cover four major areas: energy conservation and the limitation of greenhouse gas emission, preservation of areas with rich ecological diversity, protection of international waters from pollution and protection of the ozone layer.

According to the World Bank, GEF will provide practical experience in implementing environmentally sound programmes worldwide before the 1992 United Nations Conference on Environment and Development. Peter Coles

mittingly grim. George Djolov, director of the Ecology Institute of the Bulgarian Academy of Sciences in Sofia, reported that 50,000 hectares of Bulgaria near non-ferrous smelters had been so heavily polluted with heavy metals that, if US

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

International Institute for Applied Systems Analysis — seeking to secure its future.

standards were applied, all children in the region would have to be taken to hospital for detoxification.

A report from Hungary indicated that drinking water in 20–25 per cent of the country's villages is contaminated so badly with nitrates that the authorities have to bring in water in plastic bags for babies.

Of the six countries represented, Poland has made the most progress in building the necessary institutions, such as an environment ministry and credible non-governmental organizations, although the future is clouded by the recent demise of the government. The ministry is the first in Eastern Europe to outline a clear policy and to tell other ministries what they must do to fulfil environmental goals.

Poland is also the first to receive international help to clean up its environment. The World Bank in Washington has so far loaned \$18 million and the European Communities have singled out Poland and Hungary to receive environmental aid through the PHARE (Poland and Hungary Assistance to the Restructuring of Economics) programme. Together they will receive 49 million ECU (\$35.5 million).

But Western lenders are concentrating more on building up the economies of Eastern Europe than on the environment. A World Bank representative in Washington, Richard Ackermann, says that the bank is "increasingly saying that environmental issues cannot be settled until the economic issues are settled". And he adds, "The economic distortions in Eastern Europe are so enormous that it would be unwise to move in with massive investment" right away.

Although the World Bank is considering loaning as much as \$7,000 million by 1992 to all of Eastern Europe, Ackermann warns that the costs of environmental decontamination and protection are much

higher than originally thought — the United States, by comparison, spent \$640 thousand million on the environment over the past nine years.

In every country but Bulgaria, the ecological opposition movements that helped to drive the Communists from power are losing support as economic concerns come to the fore. "[The environmentalists] no longer have a common enemy", says Steve Wassersug of the Western-backed Regional Environment Centre in Budapest.

In Bulgaria, the opposition Eco-Glasnost movement continues to gather support, in part because it promises a foreign-funded clean-up. Djolov, calls this "a dangerous game" and says that the politicians "ought to know better. They're just creating illusions", he says, "and when these fail, we will have to pay the costs." He adds, "The sooner we realize that we are on our own, the better".

Those at the meeting who looked to the clean-up of the former East Germany to provide an exemplar for the rest of Eastern Europe were disappointed by the harsh picture presented by German government representatives. "The lack of functional enforcement agencies will be a major problem" in Eastern Europe, says Hans-Joachim Hermann of the Federal Environment Office in Berlin. Unlike other countries, Germany is not dependent on the worst industrial polluters for its economic well-being, says Hermann. This means it can order compliance with strict Western standards by 1996 or else shut down the plants.

In fact, Germany is counting on shutting down the worst offenders as the fastest way to reduce harmful emissions in its Eastern reaches. Of a ten per cent decrease in sulphur dioxide emissions reported so far in eastern Germany, four-fifths is a result of shutdowns, says Hermann.

A further disappointment to other countries is Germany's decision to take what it euphemistically describes as a "pragmatic" approach to toxic wastes. The state will nominally take responsibility for toxic-waste dumps in the East, absolving investors of financial responsibility for cleaning them up.

But this does not mean that the state will clean up the dumps. Instead, it is more likely just to "put a fence around" the waste sites except where they have been proved hazardous to the health of the local population, says Hermann. "We don't even have a fund to clean up toxic wastes in the western part of Germany", he says. "You can't demand that standards be higher in the East."

Steven Dickman

IIASA seeks new look

Laxenburg, Austria

THE first post-Cold-War meeting of scientific experts on the environment was recently convened by the International Institute for Applied Systems Analysis, a product of the Cold War that is now seeking to secure its future as a leading centre for environmental and economic policy analysis.

Earlier this year, the US President, George Bush, approved a renewal of official US support for IIASA, eight years

Peter de Janosi, an economist with long experience at both the Ford Foundation and the Russell Sage Foundation, hopes to build on both the scientific strengths of the institute and its rapport with Eastern European countries — several of which are members — in attracting both good researchers and more financial support. "We're not some Western institution coming in to help the natives", says de Janosi, "We *are* the natives."

One recent example of this regional solidarity was a request from the south-eastern European organization 'Pentagone' asking IIASA to provide a regional inventory of air pollution on which political decisions in the five member states (Italy, Yugoslavia, Austria, Hungary and Czechoslovakia) can be based.

IIASA will also cooperate with the Regional Environment Centre for Central and Eastern Europe in Budapest, a US-backed clearinghouse for environment information set up this past autumn.

De Janosi sees an opportunity for growth in contract research, which together with increases in dues from its member states could account for as much as a 25–30 per cent increase in IIASA's budget over his three-year term. The budget was 117 million schillings (\$12 million) in 1989.

But he admits that this path to growth is perilous for an institute that wants to be a leader in basic research. "To get the best people", says de Janosi, "we of course have to let them do what they think is important."

For example, a delicate situation arose recently when the Soviet Union, one of the largest contributors, called upon IIASA to organize and carry out a quick review by Western and Soviet economists of Soviet economic reform plans. IIASA carried out the review gladly, said de Janosi, but he does not want the institute just to "fight fires". Instead, he said, he would prefer to try to understand economic transformation of planned economies in general with a special emphasis on the environment. But a more systematic approach, including examining the economies of other Eastern European countries, will have to wait, said de Janosi, until IIASA is specifically invited by those countries. Hiring more economists is also not yet on the cards.

Despite the popularity of the topic of global warming, IIASA intends to resist pressure to address more global or North-South issues in its research. "Our work will retain its regional focus", said de Janosi, "but it will occasionally touch on global issues."

One way to expand its research base is to seek new members from Western Europe, he said.

Steven Dickman

A new paymaster for UK science?

London

THE Royal Society should be given a greater role in the allocation of research grants, to ensure that young British scientists with alpha-rated research proposals actually get funded, according to its retiring president, Lord Porter. About half of all alpha-rated proposals are now turned down by the UK research councils, Porter says.

In his farewell address to the Royal Society on 30 November, Porter envisaged

Imperial College

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Lord Porter — concerned about research councils' 'monopoly' over basic funding.

a new Science Foundation, operating at arm's length from the Royal Society, but drawing on the judgement of its thousand fellows to allocate grants. The Royal Society already distributes a small number of grants from its government grant-in-aid. This has risen by 50 per cent in real terms in the past two years, to £14 million in 1990–91, but a huge increase would be needed if Porter's aim of funding all alpha-rated research proposals is to be achieved. A new body should actually handle the money, Porter says, to preserve the Royal Society's jealously guarded independence from government.

Porter is concerned about the effective 'monopoly' that the research councils hold over basic science funding. A scientist whose highly rated research proposal is turned down by the research councils has nowhere to turn, apart from the Royal Society's own very limited programme, Porter says.

Porter would like his proposed Science Foundation to be given a new slice of government money, but accepts that the problem with British science is not just the absolute level of support, but also the way in which the money is spent.

The Science Foundation idea is Porter's own brainchild, rather than the Royal Society's. The incoming president, the mathematical physicist Sir Michael Atiyah, is so far noncommittal. Any major changes in Royal Society policy will require careful thought, he says.

Peter Aldhous

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Peter de Janosi — building on scientific strength.

after his predecessor, Ronald Reagan, cut US aid on the suspicion that the institute harboured Soviet spies. One of the conditions of the renewed US funding is that IIASA applies its skills in mathematical modelling to environmental research.

When IIASA was founded on the outskirts of Vienna in 1972, it was one of the few joint scientific initiatives between the United States and the Soviet Union. Ostensibly because IIASA may play an important role in the rebuilding of a democratic Eastern Europe, Bush has chosen to embrace the institute.

On the scientific side, IIASA has distinguished itself recently by producing the RAINS (Regional Acidification Information and Simulation) model of acidification caused by transboundary air pollution, which has been used in European negotiations. It will soon release a study on the impact of deforestation on the economy in Europe that it hopes will sway countries to do more to reduce harmful emissions.

The meeting of pollution experts demonstrates that IIASA is keen to move into the policy area, but to do this the institute will not only have to win heavy investment but also compete with an abundance of environmental and economic policy institutes in the West.

Nevertheless, the new US director,

Franz-Karl Nebuda

Tough decisions at SERC

London

SOME large projects may have to be cancelled, not just deferred, to put the UK Science and Engineering Research Council (SERC)'s troubled finances back on course, its chairman, Sir Mark Richmond, warned last week. Fresh from discussions at the Advisory Board for the Research Councils on the allocation of the 1991-92 science budget among the five research councils, Richmond said he could not see how some of the bigger components of SERC's research programme can survive.

SERC must save £40 million to avoid a shortfall in 1991-92 (see *Nature* 348, 377; 29 November 1990). No decisions have yet been taken, but Richmond said that the British involvement in two space missions, which would cost SERC more than £13 million in total, is now in question: Lyman-FUSE, a far-ultraviolet astronomy collaboration with the United States and Canada, is due for launch in 1997; Spectrum-X, an X-ray astronomy mission involving the Soviet Union and a number of the European Space Agency partners, is to be launched by 1994. Two ground-based astronomy projects already deferred, a gravity-wave observatory and an eight-metre optical telescope, will also be re-examined.

Apart from balancing SERC's books in 1991-92, Richmond wants to release more money for research grants and studentships: "I think that's the way in which really bright young scientists emerge." He became alarmed during the 1980s as grants were squeezed by the increasing costs of international programmes, and more of SERC's budget became tied up in large projects. This is partly the fault of the British public expenditure system, Richmond said, where research councils are more likely to get new money for a specific new project, rather than for support grants. "The government is not prepared to give the research councils the actual responsibility for guiding the research programme of the country", he said.

But research grants and studentships will suffer alongside major projects in SERC's immediate cost-cutting. As much as £15 million of the £40 million savings must come from studentships and grants, Richmond said.

One way to release money for smaller grants is to reduce spending on SERC's Interdisciplinary Research Centres (IRCs) over the coming years. Thirteen of these large six-year grants, which set up a team to work in a specific area of research, have been awarded over the past three years. Richmond expects fewer new IRCs in the future. The IRC is a "useful concept" in areas such as materials science

where research is centred around expensive items of equipment, although Richmond said that IRC status may have been given to some research groups that could best have been supported through smaller grants.

The large proportion of SERC's spending that is taken up by international commitments is another important problem. CERN, the Geneva-based particle physics centre, will cost SERC £9 million more than expected this year, partly because Switzerland has recently developed an inflation problem. However, Richmond acknowledged that it is not "practical politics" to negotiate new terms for participation in CERN, less than three



Richmond: large projects may have to be cancelled.

years after the United Kingdom instigated a major package of reforms to control costs.

Two other international agreements — the European Space Agency's science programme (see opposite), and the Institute Laue-Langevin (ILL) neutron source in Grenoble — are more realistic targets for savings. ILL, a collaboration between Britain, France and Germany, currently costs SERC £8.5 million a year, and SERC must soon negotiate the UK contribution over the next five years. Britain also runs the Isis neutron source at SERC's Rutherford Appleton Laboratory, although ILL and Isis are best suited to a different range of experiments. Richmond said that SERC's Science Board must look hard at the British presence at ILL.

But Richmond is determined that the soul-searching and cutbacks at SERC should not simply leave the council supporting a smaller version of its existing programme. "We must, if we possibly can, try to generate a bit of space to start something new. At the heart of it all is flexibility — you don't know when the supernova's going to go off", he said.

Peter Aldhous

ESA's science under scrutiny

London, Paris & Munich

THE European Space Agency (ESA)'s science programme, one of the Science and Engineering Research Council (SERC)'s major worries (see opposite), will be hotly debated at a two-day meeting in Paris next week. Delegates from the 13 member states hope to agree how much to spend on ESA's space science over three years from 1993 to 1995. But doubts over whether badly needed improvements to the programme's management will actually be forthcoming may delay a final decision until a high-level ministerial meeting planned for late 1991.

ESA's science budget is in trouble, plagued by cost overruns on important missions. Roger Bonnet, ESA's director of space science, has warned that some projects, such as the Huygens probe, ESA's contribution to the US Cassini mission to Saturn, may have to be cancelled.

Britain agreed the last budget increase in 1988 only after an independent review was set up to look for savings, chaired by Professor Klaus Pinkau, from the Max Planck Institute for Plasma Physics near Munich. Pinkau recommended increasing the budget to 200 million accounting units (AU) (about £140 million) in 1984 prices by 1994. But savings of 40 million AU a year are still needed for ESA to fund all the planned missions, and some new small projects.

The Pinkau report produced a list of cost-cutting options. The strict rule that awards equipment-manufacturing contracts to ESA member states in proportion to their financial contributions should be relaxed to get better value for money; and the science programme should be compensated for cost overruns forced by launch delays, and relieved from the excessive burden of overhead charges for use of ESA facilities. Missions should be costed more accurately and ESA bureaucracy reduced, Pinkau concluded.

Britain will be represented in Paris by the British National Space Centre, but the UK contribution is paid from SERC's budget. The new chairman, Sir Mark Richmond, said last week that until SERC is convinced that the Pinkau report will produce savings, "I can't see us committing ourselves".

Unlike 1988, when Britain alone resisted increases, other ESA nations are now thought to share the British concern. Delegates are reluctant to reveal their positions in advance of the meeting, but the huge costs of reunification will influence the German position. Even the French, traditionally the most enthusiastic supporters of ESA projects, are now thought to be looking for savings.

Peter Aldhous, Peter Coles
& Steven Dickman

RADIOACTIVE WASTE

Michigan faces disposal ban

Boston

EVERY week, Michigan hospitals, research laboratories and nuclear plants churn out tonnes of radioactive waste, just like those in the rest of the country. But unlike the other states, Michigan now has nowhere to dispose of this waste. And unless state officials reach a compromise with neighbouring states, some Michigan facilities may have to stop using radioactive materials, or risk drowning in their own hazardous debris.



Last month, the three states with low-level radioactive waste landfills banned all waste from Michigan until the state makes some progress with setting up its own repository. The ban was announced simultaneously by officials in Nevada, Washington and South Carolina, with immediate effect.

The move underlines the United States' quandary over low-level waste disposal: existing repositories are due to be closed and public opposition means that most states are behind schedule in opening new sites.

Some states, including California and Nebraska, have made progress in setting up new sites, but no state wants to be the first for fear that it will then become the

COMPUTER INDUSTRY

US companies join European initiative

London

THE Joint European Submicron Silicon Initiative (JESSI), a \$5,000 million collaborative semiconductor research programme set up in 1989 to help the struggling European computer industry to compete with the United States and Japan, is opening its doors to US companies. American Semiconductor Manufacturing Technology (SEMATECH), a consortium of 14 US companies, and IBM Europe, are both to join JESSI projects. Peter Aldhous

resting place of all low-level radioactive waste in the United States. Indeed, the incoming governor of Nebraska has threatened to halt progress until other states catch up.

The action against Michigan's waste illustrates the resolve of states with repositories not to accept the nation's low-level waste indefinitely. As Bill Miller, the governor of Nevada said, "this action should demonstrate to all concerned that Nevadans will not tolerate any attempts to delay closure" of the Nevada repository beyond the end of 1992.

Under the US low-level radioactive waste laws, the three states with repositories are authorized to impose sanctions against states that are not making progress towards building new low-level waste facilities. All facilities must be ready by 1993.

Michigan, designated as the repository state for a group of seven midwestern states, has just been forced to drop the last site under consideration for a low-level waste repository because of concern about nearby wetlands. It must now begin again, prospecting some 78 remaining possible areas. "We are back to the drawing board", said one Michigan official.

Michigan is the largest but not the first state to lose its rights to dump low-level radioactive wastes. Vermont, New Hampshire and Rhode Island as well as Puerto Rico and the District of Columbia have previously lost access for failure to make provisions for the wastes they generate, and other states, including Connecticut, Maine, Massachusetts, New Jersey and New York have been warned of similar action. Virtually all states are held up by battles over sites.

The federal government is trying to take a lead with several national proposals. Earlier this year, the Nuclear Regulatory Commission (NRC) announced that it would exempt "some or all regulatory controls [from] certain practices involving small quantities of radioactive material". The NRC claims that excluding low levels of radioactivity as 'below regulatory concern' (BRC) would allow it to focus its efforts on regulating activities that pose the most serious health threats.

Critics believe the move is designed to slash the cost of decommissioning nuclear power plants by allowing radioactive waste that falls 'below regulatory concern' to be disposed of in the same way as any other waste.

Although the nuclear industry has adopted the BRC framework, there are fears that a fall in the volume of low-level radioactive waste would lead to even greater disposal problems by creating an economic disincentive to build any new repositories.

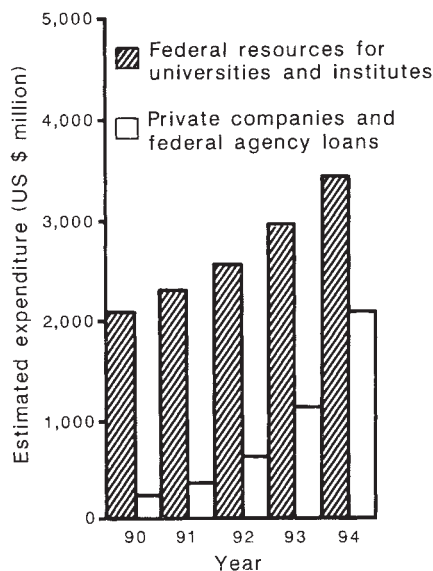
Seth Shulman

BRAZIL

More money for applied research

São Paulo

BRAZIL intends to more than double the amount it spends on research and development by 1994, the end of President Fernando Collor de Mello's term, Secretary of Science and Technology José Goldemberg said last week. At present, just 0.7 per cent of the country's gross national product is spent on research and development (R&D), but Goldemberg said more is needed to make Brazilian industry competitive in the international market. Private Brazilian companies currently spend US\$200 million a year on R&D, and receive another US\$60 million from federal agencies. Goldemberg wants to increase the federal contribution to US\$1,430 million in 1994 from these agencies — an increase of 2,283 per cent — and to encourage industry to spend an additional US\$474 million of its own money. Resources



Brazil's R&D expenditure plans

for universities and research institutes would grow from US\$2,070 million this year to about US\$3,462 million in 1994 under his plan.

The president of the Brazilian Society for the Progress of Science, Ennio Candotti, is sceptical. "Before thinking of 1993 or 1994, we must see how scientific institutions will survive the penalty they face in 1990 and 1991", he says.

■ DESPITE the objections of their armed forces, Argentina and Brazil agreed last week that neither would build nuclear weapons. The agreement, signed by Argentinian President Carlos Saul Menem and his Brazilian counterpart, Fernando Collor de Mello, is intended to convey that both are "responsible" nations in nuclear matters, said Brazil's Secretary of Science and Technology, José Goldemberg.

Ricardo Bonalume

School science faces reform

Washington

THIRTY years after Sputnik launched the space age and the short-lived renaissance of US science education, government officials are hoping to regain those glory days with a radical rethinking of education. In the first set of major reforms since the 1950s, federal policy-makers are preparing to set a national agenda for science instruction in the pre-college years, led by a new presidential education initiative to be announced early next year.

Joining the government in the battle to save US science education, two national organizations are preparing programmes that would have the effect of totally restructuring science teaching on a national scale — a substantial change from the regional approach that now dictates how most US students learn science.

Since the mid-1950s, US science and mathematics education has declined in virtually every category. Searching for an explanation, policy-makers have pointed to teachers who have lost touch with the research world, teaching methods that date back decades and a general resistance to new educational materials. A series of international comparisons each more devastating than the last, has fuelled the debate for educational reform.

Thomas Ratchford, associate director for policy and international affairs in the White House Office of Science and Technology Policy acknowledges that “at the K-12 [between the ages 5 to 17 or 18] level, we are in deep trouble”, and says that reform is a top priority for this administration. In a speech raising the issue last year, President Bush pledged that, “by the year 2000, US students will be the first in the world in science and mathematics achievement”.

One of the major obstacles to reform is the sheer size of the US education system, which serves 46 million children in 80,000 schools within 16,000 school districts. The decentralization of authority and allocation of resources make the capacity for rapid change impossible. The responsibility for education rests primarily with the state legislators and the local school boards, and there is no intention on the part of the administration to change that situation, says Ratchford.

Although the federal role in education is relatively small — less than 6 per cent of the \$200 thousand million that is spent annually on education — Ratchford believes that injecting more money into science education is not the answer. Investment in pre-college education has increased in real terms over the past ten years and “we have very little, if anything, to show for it in terms of test scores”.

In order to achieve maximum leverage in science education with limited resour-

ces, the administration has set up an Education and Human Resources Committee within the Federal Coordinating Council on Science, Education and Technology (FCCSET), a body led by the president's science adviser and made up of top officials from 11 science-related agencies (see *Nature* 348, 97; 8 November 1990). The committee will address policy issues that cross agency lines, identifying gaps, promoting increased agency co-operation and eliminating unnecessary duplication. The primary focus of the interagency effort will be pre-college education, which is a priority in the president's national goals for education. Specific targets of the science-education initiative will be announced in the president's January budget request for 1992.

The United States has no nationally mandated curriculum for science, although most states require students to study science of some sort for a minimum of two years in order to graduate from high school. In general, the United States adopts the ‘layer-cake’ approach to science teaching, where each subject is taught as a condensed year-long course with no attempt being made to integrate the disciplines. Students take biology one year, chemistry the next, physics the year after. More than 50 per cent of students opt out of science at the first opportunity, and only 19 per cent of high-school graduates have studied physics as a single subject; 40 per cent take chemistry.

NSTA's Scope, Sequence and Coordination (SS&C) Project and Project 2061, run by the American Association for the Advancement of Science (AAAS), offer

two different approaches to a parallel problem — the total restructuring of science teaching on a national scale. The SS&C initiative, which is funded by the National Science Foundation (NSF) and the Department of Education (DoEd), calls for a scrapping of the layer-cake approach and recommends that all students learn biology, chemistry, physics and Earth/space science each year for six years in grades 7–12. Experimental trials with SS&C began in September in selected schools at five sites, including the state of California and the Houston Independent School District.

By contrast, Project 2061 is a three-phase 25-year effort designed to raise the scientific literacy of all Americans, with the expectation that this would provide “a greater pool from which to entice future scientists”, says Jo Ellen Roseman, curriculum director of Project 2061. *Science for All Americans*, the product of phase I, re-examines the education process from kindergarten right through to high school and establishes learning goals in science, mathematics and technology. A variety of curriculum models should be available by spring 1993.

The United States may be in a new era of educational reform, but being able to turn the situation around is far beyond the ability of any one single sector to influence. Nationwide reform will need to encompass all aspects of the education system. If the focus is only on teachers and a school's organization or its curriculum do not change, those teachers will “have brilliant ideas and no place to take them”, says David Florio, director of policy studies in DoEd's Office of Evaluation and Research.

Diane Gershon

Massive change needed to halt decline

Washington

Six months after he was fired from the top National Science Foundation (NSF) education post for his outspoken criticism of the US educational efforts, Bassam Shakhashiri is back in the classroom. Now a chemistry professor at the University of Wisconsin, Shakhashiri argues that simply restructuring federal support of US science education will not be enough to halt the country's precipitous educational decline.

Six years ago, when Shakhashiri joined NSF during the Reagan administration, educational activities at the agency had been all but shut down. But during the next half-decade, Shakhashiri's single-minded effort to raise the profile of NSF in science and mathematics education helped to boost the annual budget from \$50 million to more than \$300 million.

Yet even that is not enough, says

Shakhashiri, who feels that “what is needed is \$600 million, so that NSF can exert its leadership role among the scientific, educational and business communities . . . to bring about statewide systemic change”.

Shakhashiri makes it clear that he is not interested in incremental change. The situation that the United States now faces “is, by far, more serious and more consequential than that faced in the immediate post-Sputnik era”, he says.

Then, reform efforts in the United States concentrated on regaining pre-eminence in the fields of science, mathematics and engineering. In what Shakhashiri describes as a twin mission, he acknowledges the need for continued support of the best and the brightest students, but is emphatic about the need to communicate science to the non-scientists.

Diane Gershon

Consciousness and time

Ian M. Glynn

Roger Penrose has suggested that when we consider consciousness the usual physical rules for time may not apply. But that notion is based on a false interpretation of physiological observations.

IN the tenth and last chapter of his book *The Emperor's New Mind*¹, the mathematician and theoretical physicist Roger Penrose warns that "we may actually be going badly wrong when we apply the usual physical rules for *time* when we consider consciousness!". That alarming statement — the adjective is Penrose's — is based largely on the results of a series of experiments by the neurologist Benjamin Libet and his colleagues²⁻⁵. The same results had earlier impressed the neurophysiologist John Eccles, who, in *The Self and its Brain*⁶, a book that he wrote with Karl Popper, says of them that they do "not seem to be explicable by any neurophysiological process". If Penrose and Eccles are right, both their conclusions and the results of Libet *et al.* are of fundamental importance. In this article however I shall argue that the interpretation of their own experiments by Libet *et al.* is flawed, and that even if that interpretation is sound it does not imply anything mysterious about the relation between consciousness and time.

Sensory experience

The experiments at issue purport to show that when, in response to a stimulus, a subject has a conscious sensory experience, that experience seems to the subject to occur not at the time when the necessary neuronal events are complete (as might be expected), but at the much earlier time when it first becomes possible to detect an electrical response in the cerebral cortex (early positive evoked potential). Because, in the experiments of Libet *et al.*, the early positive evoked potential occurred only 15 milliseconds (ms) after the stimulus, whereas several hundred milliseconds had to elapse before the neuronal events necessary for the sensation to occur were complete, it is claimed that there must have been "an automatic subjective referral of the conscious experience backwards in time" of the order of several hundred milliseconds. Such behaviour is less bizarre than that of Lewis Carroll's White Queen, who applied sticking plaster and screamed *before* she pricked her finger, but it is remarkable enough. Yet unlike Alice, who was sure that there was something wrong with the White Queen's arguments, Eccles and Penrose accept those of Libet and build on them. One suspects that Alice, given access to volume 102 of the journal *Brain* (ref. 4), would have been more critical —

and she would have been right to be.

Libet and his colleagues point out, correctly, that there is no method for determining the absolute timing — the 'clock time' — of a subjective experience; all that can be done is to compare the subjective timing of different experiences and make use of any constraints about timing that are known to exist. In their experiments, they therefore compared the subjective timing of the conscious sensations in the skin caused by three kinds of electrical stimulation — direct stimulation of the skin; stimulation of the somatosensory cortex (the bit of cortex that is the primary receiving area for skin sensation); and stimulation of the medial lemniscus (a tract carrying sensory information through the brainstem to the thalamus, whence it is relayed to the somatosensory cortex). For stimulating the skin, single shocks were sufficient to cause a conscious sensation, though in some experiments stimulation was by short trains of shocks each too weak to be effective singly; to cause a conscious sensation by stimulating the somatosensory cortex or the medial lemniscus, it was necessary to give a train of shocks — say 60 shocks per second for several hundred milliseconds. The duration of the shortest train that is effective in causing a sensation varies with the strength of shock. This duration Libet calls the minimum or 'utilization' train duration, and it ranged between 200 and 500 ms in the different experiments. All of the shocks were very weak, generally only a little above the minimum necessary to cause a response. The experiments were carried out (with informed consent) on patients who had had relevant portions of their brains exposed for therapeutic reasons that we need not discuss here.

There are three sets of experiments to consider. The first³ is concerned with determining the interval between giving a single electric shock to the skin of a subject and the subject's awareness of the resulting sensation. The second⁴ compares the timing of the sensation elicited by stimulating the skin directly, with the timing of the sensation elicited in the same bit of skin by stimulating the appropriate area of the somatosensory cortex. The third⁵ compares the timing of the sensation elicited by stimulating the skin directly, with the timing of the sensation elicited by stimulating the medial lemniscus.

In the first set of experiments, Libet and his colleagues³ found that when they

stimulated the somatosensory cortex with a train of weak shocks of known strength, at a frequency of 60 shocks per second, no sensation was felt in the skin unless the stimulation was continued for at least 300–500 ms. This showed that the sensation could not have occurred earlier than 300–500 ms after the beginning of the stimulation. The sensation caused by stimulating the cortex was felt in a particular area of skin, and this area could also be stimulated directly, a single shock sufficing to cause a sensation. (The sensation caused by direct stimulation differed slightly in quality from that caused by cortical stimulation.)

Double stimulus

What happened, Libet and his colleagues asked, if both kinds of stimulation were used? It turned out that if the single shock to the skin was given during the cortical stimulation, only the sensation associated with the cortical stimulation was experienced; somehow the effect of the skin shock was blocked. More remarkably, the effect of the skin shock was blocked even if the skin shock was given 300–500 ms *before* the beginning of the cortical stimulation. They argued that this 'retroactive masking' proves that the neuronal events necessary for conscious awareness of the skin stimulation must take at least 300–500 ms. Furthermore, in a later paper⁷, Libet points out that, to achieve retroactive masking, it was necessary to stimulate the cortex for at least 100 ms, so the neuronal events responsible for the skin sensation cannot have been completed until at least 100 ms after the beginning of cortical stimulation; the estimate of the minimum time taken for the neuronal events necessary for conscious awareness of stimulation of the skin should therefore be increased to 400–600 ms. These long durations may sound improbable and contrary to everyday experience, but, because the stimuli used were very weak and quite unphysiological, everyday experience is no guide. Reaction times can, of course, be much shorter than 400 ms, but Libet⁷ points out that one can react to an event without being consciously aware of it.

An alternative explanation of the experiments that seem to demonstrate retroactive masking has been proposed by Churchland⁸. She suggests that the cortical stimulation might not prevent the conscious sensation that normally follows stimulation of the skin, but might merely

erase the memory of it, so that it was not reported. Libet⁷ rejects this interpretation, using the strange argument that, in some subjects, cortical stimulation caused retroactive enhancement rather than retroactive masking. But although Churchland's hypothesis cannot be excluded, in the absence of any positive evidence for it I shall assume that Libet's interpretation of the retroactive masking experiments is correct.

In the second set of experiments, Libet *et al.*⁴ again looked at the effects of combining cortical stimulation and skin stimulation, but this time the skin that was stimulated was on the opposite side of the body to the skin to which the sensory response to cortical stimulation was referred. So in these experiments there was no masking. Stimulation of the cortex made the subject feel a sensation in a particular bit of skin on one side; stimulation of the skin made him feel a sensation — a rather different sensation — in the corresponding bit of skin on the other side.

What Libet and his colleagues did was to vary the timing of the two stimuli — the shock to the skin and the series of shocks to the cortex — and ask the patient which sensation started first, that on the right or that on the left. (In control experiments, the stimulus on both sides was a shock to the skin.) Because the experiments on retroactive masking implied that the single shock to the skin led to a sensation only 400–600 ms later, the authors expected that when the shock to the skin was given, say, 200 ms after the start of the train of shocks to the cortex, the sensation caused by the skin shock would begin later than the sensation caused by the cortical shocks (see *a* in Fig. 1). In fact they found that the sensation caused by the skin shock was felt first. From this they argue that the subjective timing of the sensation resulting from direct stimulation of the skin is referred back to a time shortly after the stimulus and long before the completion of the neuronal events that are necessary for the sensation to occur. Furthermore, because direct stimulation of the skin, unlike cortical stimulation², gives a surface-positive evoked potential in the cortex after only 15 ms, they suggest that this early evoked potential might be the time marker to which the sensation is later referred back.

This notion that the early positive evoked potential serves as a time marker is supported by the third set of experiments⁴, in which the results of direct stimulation were compared with the results of stimulation of the medial lemniscus. As with cortical stimulation, generation of a conscious sensation by stimulation of the medial lemniscus requires a train of shocks lasting a few hundred milliseconds, the exact time depending on the strength of shock. Unlike cortical stimula-

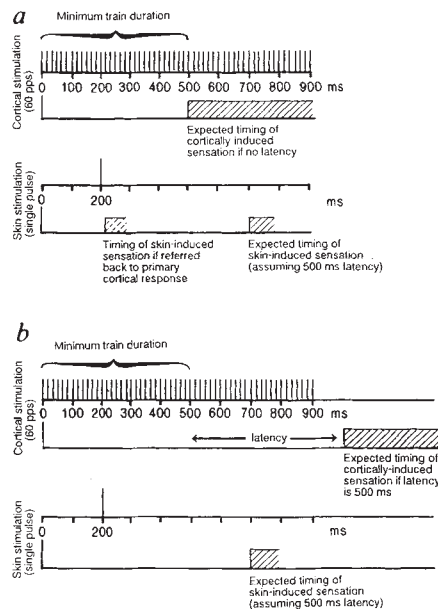


Fig. 1 *a*, Illustration of the second type of experiment, reported by Libet *et al.* in ref. 4 (modified from Fig. 1 of that paper). A train of shocks was given to the somatosensory cortex, and 200 ms after the start of the train a single shock was given to the skin on the opposite side from that to which the cortically induced sensation was referred. The minimum train duration is the duration of the shortest train of cortical shocks capable of causing a sensation, and Libet *et al.* assume that the sensation starts as soon as the minimum train duration is reached. To explain why the subject reported that the skin-induced sensation preceded the cortically induced sensation, Libet *et al.* postulate that the skin-induced sensation is subjectively referred back to the time of the primary cortical response, which occurs about 15 ms after the stimulus. *b*, Demonstration of how a latency of a few hundred milliseconds following the end of the minimum train duration makes it unnecessary to postulate back-referral. In both *a* and *b* the latency of the skin-induced sensation was assumed to be 500 ms (based on experiments showing retroactive masking).

tion, however, stimulation of the medial lemniscus does cause an early positive evoked potential in the cortex similar to that seen when the skin is stimulated directly. And it turned out that, when the shock to the skin was made to coincide with the start of the train of shocks to the medial lemniscus, the sensations tended to begin at about the same time. This, of course, is what would be expected if each sensation were referred back to the time of the corresponding early positive evoked potential.

An obvious weakness in the argument from the second set of experiments, first pointed out by Churchland⁴, is the implicit assumption that, with cortical stimulation, there is no significant latency — that is, the conscious sensation is felt as soon as (or very shortly after) the train duration reaches the minimum or 'utilization' train duration (*a* in Fig. 1). For if there were a substantial interval between the comple-

tion of the minimum train duration and the moment at which the cortically induced sensation was felt (*b* in Fig. 1), there would be no reason to expect the sensation caused by stimulating the skin to follow the sensation caused by stimulating the cortex; and no need to postulate "automatic subjective referral of the conscious experience backwards in time". The implicit assumption that there is no significant latency was made despite the explicit statement (page 199 of Libet *et al.*) "that the conscious experience for a liminal cortical stimulus could not begin before the end of the utilization train duration, though it might begin afterwards" (my italics).

In a later paper, however, Libet⁷, replying to Churchland's criticism, claims that the extent to which the skin stimulus needed to be delayed, for the skin-induced sensation and the cortically induced sensation to start simultaneously, was approximately equal to the minimum train duration; and that this was true whether the latter was set at 500 ms (with very weak shocks) or at 200 ms (with stronger shocks). And he argues that "the simplest and best interpretation of [this] consistent experimental relationship" is that the skin-induced sensation is referred backwards in time to a point almost immediately after the stimulus, and the cortically induced sensation is felt as soon as the minimum train duration is complete.

This defence of the assumption that there is no significant latency is unsatisfactory in two ways. First, the agreement between the minimum cortical train duration and the period by which the skin stimulus needed to be delayed to achieve synchrony is not impressive. Two groups of experiments are discussed in ref. 4. The first group (page 200), which involved only a limited number of tests on each subject, was not reported in detail, but the agreement was good: minimum train duration 500 ms; estimated delay for synchrony 400–500 ms. Of the three experiments reported in detail (Table 3b of ref. 4), however, only the first (subject J.W.) gave good agreement: minimum train duration 200 ms; estimated delay for synchrony 220 ms. In the second (subject C.J.), the scatter was too great for useful conclusions to be drawn. In the third (subject M.T.), the minimum train duration was 200–300 ms, but the estimated delay for synchrony was only 95 ms. Libet *et al.* dismiss this last discrepancy on the grounds that, when the sensation was cortically induced, M.T. "reported having a 'hunch', or preconscious type of feeling. . . that 'something' was building up before the instant at which he himself felt that he was actually aware of a somatic sensation . . ."; this would tend to make his reported timings for onset [of the cortically induced sensation] earlier than warranted by the criterion of subjective

experience". To accept this explanation is to cast doubt on the validity of all the reports of timings of cortically induced sensations.

The second way in which Libet's defence of the assumption of no significant latency is unsatisfactory, is that, even if one accepts the equality between the minimum train duration and the delay necessary to achieve synchrony, it does not follow that the latency is negligible. Libet⁷ is mistaken in suggesting that, to be compatible with the observed equality, the duration of a latency that was not negligible would have to be strictly related to the minimum train duration in a way that there was no physiological reason to expect. A latency that was independent of the minimum train duration, and that happened to be similar to the latency of the skin-induced sensation, would account for the equality over the whole range of minimum train durations. And, for the cortex-induced sensation and the skin-induced sensation to have similar latencies seems at least as likely as that the first should be negligibly small when the second is supposed to be several hundred milliseconds.

The finding, in the third set of experiments, that the sensation induced by stimulating the medial lemniscus, and the sensation induced by stimulating the skin, are felt at about the same time if the stimuli start at the same time, is certainly predicted by the theory of Libet *et al.*; but it does not prove that theory. For the finding can also be explained by supposing that the sum of the minimum train duration for stimulation of the medial lemniscus and the latency of the medial-lemniscus-induced sensation is roughly equal to the sum of the minimum train duration for stimulation of the skin (which in these experiments, in an attempt to match quality of sensation, was achieved by a brief train of individually subthreshold shocks) and the latency of the skin-induced sensation. A more stringent test of the theory of Libet *et al.* would be to see whether the skin-induced sensation and the medial-lemniscus-induced sensation remained synchronous (for stimuli starting together) when the minimum train duration for stimulation of the medial lemniscus was varied over a wide range by changing the strength of shock. That experiment seems not to have been done; in the experiments summarized in Table 2 of ref. 4 the minimum train duration was always in the range 200–300 ms.

The arguments given here show that the evidence for the "automatic subjective referral of the conscious experience backwards in time" is unconvincing. But suppose for a moment that Libet and his colleagues are correct in claiming that such backward referral does occur — in the sense that, when someone is subjected to two stimuli, the order of onset of the



Fig. 2 The birth of Tamar's twins — an early use of a time marker. The twin labelled with the scarlet thread (not shown) is being born; the other twin is just distinguishable lying on the floor. The figures on the left illustrate an earlier stage in the story, in which Tamar proved Judah's paternity by showing the pledges she received from him in return for sexual favours. (From the twelfth century, Greek, Seraglio Octateuch, Istanbul.)

elicited sensations is determined by the relative timing of the early positive evoked potentials, even when that order differs from the order predicted from indirect estimates of the time required for completion of the necessary preparatory neuronal events (what Libet and his colleagues refer to as "the achievement of neuronal adequacy"). How significant or surprising would this be?

It is important to note that no sensation can actually begin before the necessary preparatory neuronal events are complete. That is not a matter of physiology but a logical consequence of the meaning of the words. Unless we postulate the existence of mysterious non-neuronal events, it also follows from the meaning of the words that any sensation must begin as soon as the preparatory neuronal events are complete (though presumably further neuronal events accompany the sensation). But the time taken for the completion of the preparatory neuronal events may vary with the circumstances, and, in particular, with what else is going on in the cortex. It follows that the interval between a stimulus and the onset of the sensation it elicits may not be constant. Furthermore, the conscious awareness of sensations of different kinds must involve different neural pathways with different delays.

For both of these reasons, unless there is some mechanism that can delay or accelerate the completion of the preparatory neuronal events necessary for a conscious sensation, the order in which the events necessary for each of two sensations are completed may not be the order in which the relevant stimuli occurred. In analysing complex situations it may be important to become conscious simultaneously of things that happen at the same time. It is therefore conceivable that a mechanism

has evolved for achieving this synchronization. There is no evidence that such a mechanism exists; but if it does, one way in which it might work would be to set up time markers early on in the sequence of preparatory neuronal events associated with each conscious sensation, and then, referring back to these time markers, to arrange the sensations in proper order by delaying or accelerating the completion of each sequence. A mechanism of this kind would not seem to require anything that would justify Penrose's warning.

The idea of using time markers to establish precedence between events of some duration, when what matters is the order in which the events start, is not new (Fig. 2). When Tamar, who had been made pregnant by Judah when he mistook her for a harlot, was giving birth to Judah's twin sons (*Genesis* 38), one twin put his hand out first and the midwife bound his wrist with a scarlet thread to secure his rights as first born; the hand was then withdrawn and the other twin emerged first. If Libet and his colleagues are correct, the early positive evoked potentials are the brain's scarlet threads. But Tamar's midwife was not tampering with time. □

Ian M. Glynn is in the Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, UK.

1. Penrose, R. *The Emperor's New Mind* (Oxford Univ. Press, 1989).
2. Libet, B., Alberts, W. W., Wright, E. W. Jr & Feinstein, B. *Science* **158**, 1597–1600 (1967).
3. Libet, B. in *Handbook of Sensory Physiology* Vol. 2 (ed. Iggo, A.) 744–790 (Springer, New York, 1973).
4. Libet, B., Wright, E. W. Jr, Feinstein, B. & Pearl, D. K. *Brain* **102**, 193–224 (1979).
5. Libet, B. in *Models of Brain Function* (ed. Cotterill, M. J.) 35–49 (Cambridge Univ. Press, 1990).
6. Popper, K. R. & Eccles, J. C. *The Self and its Brain* 256–259, 362–365 (Routledge, London, 1983).
7. Libet, B. *Philosophy Sci.* **48**, 182–197 (1981).
8. Churchland, P. S. *Philosophy Sci.* **48**, 165–181 (1981).

The metre and the pendulum

SIR—In his Commentary on 'The metre and the pendulum' (*Nature* **348**, 105; 1990), J. H. Freeman remarks that no records appear to exist on contacts that might actually have taken place between France and England following the Act of the French National Assembly of 8 May 1790. In this act, as well as proposing the establishment of a new system of weights and measures, the Assembly instructed Louis XVI to write to the King of England asking for the cooperation of the British Parliament in the venture.

In fact records of such contacts do exist and were published in *Nature* (**69**, 425–427; 1904) by Herbert McLeod. Briefly, the events, as recounted in some detail by McLeod were as follows:

5 February 1790: Sir John Riggs-Miller (1744–98, Member of Parliament for Newport 1784–90) speaks to the House of Commons on the subject of uniformity of units of measurement. He makes various proposals, all of which are agreed to, for a series of inquiries into present practice among the towns and shires of England.

Not the first

SIR—In a recent Product Review article (*Nature* **345**, 747; 1990), I. Jardine writes that "electrospray ionization has been shown to be useful for the molecular weight and structural analysis of carbohydrates and . . ." (ref. 4: R. D. Smith *et al. Analyt. Chem.* **62**, 882; 1990) and then "When used with high-resolution mass spectrometers . . . ESI should, in future, allow complex mixture analysis . . .".

I must say that the first application of ESI/AP/MS (Russian abbreviations MS/ERIAD) "for the molecular weight and structural analysis of carbohydrates" has been demonstrated by us in our work in *Bioorganicheskaya Khimiya* in 1986 as was indicated in the review by R. D. Smith *et al.* (ref. 26). In the following works (P. W. Bezukladnikov, L. A. Elyakova & O. A. Mirgorodskaya *Bioorg. Khim.* **10**, 1318–1325; 1989; P. W. Bezukladnikov, L. A. Elyakova, T. N. Zvyagintseva & O. A. Mirgorodskaya *Khim. Prirod. Soedin.* **1**, 54–59; 1989; P. W. Bezukladnikov & L. A. Elyakova, *Carbohydr. Res.*: in the press) we showed that MS/ERIAD "allows complex mixture analysis" of carbohydrates not "in the future", but now on the instrument created by the Institute of Analytical Instrumentation (Leningrad).

P. W. BEZUKLADNIKOV
Pacific Institute of Bioorganic
Chemistry,
Far East Branch of the Academy of
Sciences of the USSR,
Vladivostok 69002, USSR

He suggests the pendulum as the basis for a new unit of length.

28 March 1790: Talleyrand (Bishop of Autun) writes to Riggs-Miller congratulating him on his initiative and says that he is about to make a similar proposal to the French National Assembly.

1 April 1790: The House of Commons establishes a committee (which includes Riggs-Miller) to examine and report upon weights and measures.

13 April 1790: Another speech in the House by Riggs-Miller announcing his contacts with Talleyrand and going into further detail on proposals for new units and standards.

8 May 1790: Decree adopted by the French National Assembly proposing a new system for weights and measures and instructing Louis XVI to write to the King of England.

22 May 1790: Letter to the British Foreign Secretary from the French Ambassador in London formally inviting the House of Commons to join the National Assembly in establishing a new system of weights and measures. The Ambassador proposes that the Royal Society and the Académie des Sciences jointly take up the matter.

11 June 1790: Parliament is dissolved. At the subsequent election, Riggs-Miller loses his seat at Newport. In the new parliament, the committee on weights and measures is not reappointed.

3 December 1790: The Foreign Secretary writes to the French Ambassador saying that although the House of Commons has discussed his proposal, no formal motion has been put and therefore nothing can be done. In any case he adds "the whole scheme looks impractical".

And there the matter rested.

T. J. QUINN

Bureau International des Poids et
Mesures,
Pavillon de Breteuil,
F-92312 Sèvres Cedex,
France

National citations

SIR—In a recent review article in a US journal (*Science* **246**, 465; 1989), Konishi *et al.* argue convincingly for the usefulness of birds as study subjects, but their citations of the literature show a strong US bias suggesting that US ornithologists have made a very large contribution to the progress of avian biology.

Scientists based in the United States had written 77 per cent of the 214 papers cited in the review. The preponderance of US papers was more marked for those published in the 1970s and 1980s (91 per cent of 51 and 85 per cent of 105 respectively) than for those published before 1970 (51 per cent of 59).

By way of control, and following the statement of Konishi *et al.* that the 1985 volume of the *Zoological Record* contains more than 9,300 articles on birds, I picked at random 25 pages from each of the 1965, 1975 and 1985 volumes to estimate the US contribution at those times. The proportions of US papers were 21 per cent (of 930 papers) in 1965, 19 per cent (of 770 papers) in 1975 and 19 per cent (of 668 papers) in 1985.

Because English-speaking scientists are notorious for ignoring non-English publications, I have also taken from my sample of articles cited in the *Zoological Record* only those written in English. I find that the proportion of papers by US-based scientists was 37 per cent (of 523 papers) in 1965, 33 per cent (of 440) in 1975 and 28 per cent (of 460) in 1985. Thus the over-representation of US citations in Konishi *et al.* does not stem from a language problem. Indeed, the data suggest that there should have been an increase in the citation of non-US papers during this period, for the proportion of papers written in English increased from 56 per cent (of 930) in 1965 to 69 per cent (of 668) in 1985.

The over-representation of US citations in US papers has been remarked on previously (P.-H. Enckell *Nord. Ecol. Newslett.* **39**, 1; 1988), in a comparison of the citation practices of US and Swedish ecologists.

What is the explanation of this bias? There are four possibilities, not necessarily mutually exclusive, as follows: (1) non-US scientists publish relatively inferior science, so that their papers are not worth citing; (2) many US studies are duplicated elsewhere, in which case it is easier to cite the more readily available publications; (3) US scientists do not see European journals, so that non-US studies are simply overlooked; and (4) it is a consequence of chauvinism.

I do not believe the first explanation is justified, but some combination of the other three may account for the data I have quoted. The question can be resolved only by further detailed study of citation practices.

The result is that US-based scientists get a fair share of European citations plus a disproportionate share of the US citations. This bias in citation frequencies may have important consequences: the dissemination of scientific ideas is slowed down, as the same hypotheses are re-invented time after time. And, at an individual level, European and US scientists competing for the same job may turn up with different citation scores, not because their research differs in quality, but because they happen to live on different continents.

ANDERS PAPE MØLLER

Department of Zoology,
Uppsala University, Box 561,
S-751 22 Uppsala, Sweden

Photonic band-gaps bite the dust

Hopes that dielectric materials in which the transmission of certain frequencies would be forbidden seem to have been disappointed by the difficulty of realizing expectation and, now, by calculation.

IF only it were possible to make dielectric materials in which electromagnetic waves cannot propagate at certain frequencies, all kinds of almost-magical things would be possible. For example, it would be possible to embed in such a material atoms or molecules in excited states that would literally be incapable of getting rid of their excess energy by radiation if its frequency were in the forbidden range; the result might be a novel state of matter in which atoms or molecules are forever bound to the photons they would emit if their surroundings would accommodate them.

The goal should not be beyond reach. There are familiar examples of how the propagation of waves with certain frequencies is denied. Waveguides, for example, will not accept or transmit microwaves with a frequency below a cut-off frequency determined by their dimensions. In metals, there are allowed and forbidden energies, proxies for frequencies, at which electron waves can propagate without exponential attenuation. The most relevant analogy is that of the perfectly reflecting mirror constructed from several alternating layers of dielectric with alternately high and low refractive index, each of which is a quarter-wavelength thick. So why not look for a three-dimensional dielectric solid in which certain optical frequencies would be forbidden in all directions?

Only a year ago, this esoteric search seemed to have succeeded. But now it has been disappointed by the killjoy calculations of two groups of theoreticians. But the non-existence theorem now offered is not conclusive while, for the rest of us, the brief history of this affair is entrancing.

The story begins with Eli Yablinovitch, at the Navesink Research Center of Bell Communications Research, who drew attention to the potential value of materials in which the transmission of light of certain frequencies would be literally impossible. To the extent that the efficiency of devices as different as solar cells and solid-state lasers are determined by the rate at which electrons and holes recombine (with the emission of radiation) in semiconductors, it would plainly be of some importance if that were naturally suppressed by the structure of the material (*Phys. Rev. Lett.* **58**, 2058; 1987).

The potential of the technique was widely and quickly applauded. That some electronic devices might be made dramatically more efficient pointed to an evi-

dently important prize. More surprising was the prediction that the properties of many molecular systems, when embedded in such materials, would be modified radically; disassociation energies of homonuclear diatomic molecules, for example, might be substantially reduced. The explanation is that interatomic interactions entailing the interatomic transfer of even virtual photons in the forbidden ranges of frequency will no longer occur.

But how to design and then construct a dielectric material in which the transmission of photons of certain frequencies is forbidden, whatever the direction in which they travel? Photonic band-gap is the word. Three years ago, Yablinovitch outlined a recipe, a three-dimensional analogue of the perfectly reflecting dielectric sandwich, for constructing such a solid from dielectric materials with sharply different refractive indices. If λ is the wavelength of the light whose propagation is to be suppressed, first lay down a layer of one material with optical thickness $\lambda/4$, etch in that a chequerboard array of square pits with the same dimensions, cover that with a similar layer of the second dielectric in which should be etched a similar chequerboard pattern of holes interdigitating the first . . . and then repeat the process indefinitely. The end result is a three-dimensional array of blobs of dielectric material with different refractive indices, arranged with the symmetry of a face-centred cubic lattice.

Yablinovitch himself seems to have spent much of the interval since 1987 trying to turn his recipe into reality. To avoid the cost of turning an epitaxial semiconductor fabricator into an experimental tool, he worked with microwaves, meaning that suitable gap-ridden dielectric solids might be fabricated from bulk pieces of dielectric with different refractive indices (dielectric constants) using precision machine-tools. It seems to have been something of a struggle. Last year, Yablinovitch and T. G. Gmitter reported the outcome of their "tedious cut-and-try", or "empirical Edisonian" approach to the problem of realization (*Phys. Rev. Lett.* **63**, 1,950; 1989). "Dozens" of structures were machined out of low-loss dielectric material, some containing as many as 8,000 "atoms", or blobs, material with contrasting dielectric constants. Some "atoms" were dielectric spheres, others were spherical hollows filled with air in a sheet of dielectric material.

Most of the "crystals" made in this way

appear not to have inhibited microwave transmission, but Yablinovitch was able to report that, provided that the two refractive indices of the dielectric components stand in the ratio of at least 3.5 to 1, a photonic band-gap does indeed appear. The most successful of his crystals consisted of spheres filled with air machined into a proprietary dielectric material with a refractive index of 3.5 in the frequency range between 1 and 20 GHz. Only 14 per cent of the crystal volume was made of dielectric, so that the structure was more closely packed than expected for a face-centred close-packed structure.

In the most successful structure, there seemed to be a band-gap, the width of which seems to have been dependent on the direction of propagation and to be a few per cent of the central frequency whose transmission is inhibited. One evident disappointment was that the contrast of refractive indices required for inhibition turned out to be much greater than expected; another is that there seemed to be no simple relationship between the structure of the composite dielectric crystals and the frequencies at which the band-gaps appear.

This is where the killjoys come in. Those concerned are K. M. Leung and Y. F. Liu from the Brooklyn Polytechnic University, New York, and Ze Zhang and Sashi Satpathi from the University of Missouri at Columbia (*Phys. Rev. Lett.* **65**, 2646 and 2650 respectively; 1990). Each has carried through a calculation of the behaviour expected of a two-component dielectric lattice when full account is taken of the vector character of electromagnetic radiation. It is simply a matter of eliminating the magnetic fields from Maxwell's equations and of treating separately the propagation of plane waves with various directions of travel. There emerges an equation in the form of a determinant of infinite dimension that can be solved numerically by truncation.

What emerges (most clearly from the work of Zhang and Satpathy) is that true band-gaps cannot appear, but that there are circumstances in which the density of transmission states as a function of frequency is not a smoothly varying function, but rather one in which there are so few transmission states in a certain frequency range that the region concerned can be called a "pseudo-gap". On past form, Yablinovitch will no doubt set out to build just such a microwave crystal. That is as it should be.

John Maddox

On to the second generation

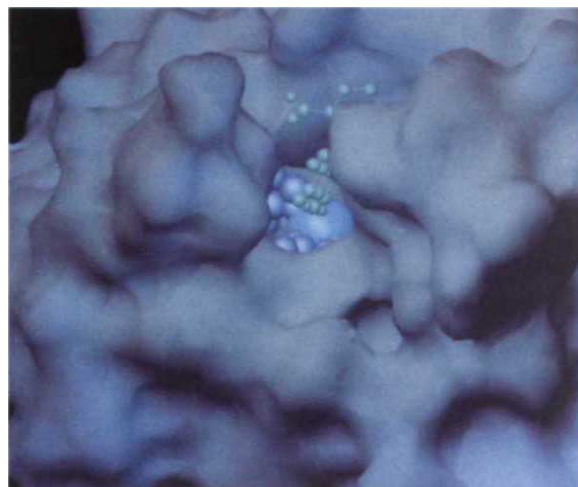
Robin J. Leatherbarrow

It is only four years since catalytic antibodies were first produced^{1,2}, yet at a recent meeting* the talk was of the latest second-generation molecules that promise increased efficiency, greater applicability and increased ease of generation. Although designer catalysts cannot yet be produced on demand, the principles are now better understood and it will surely not be too long before tailored syntheses that require stereochemical control can be attempted³.

Current ideas on enzyme catalysis stem from Linus Pauling (reviewed in ref. 4). Enzymes catalyse reactions by binding, and hence stabilizing, the transition state of the reaction in preference to the ground state. Antibodies are extremely efficient at binding, but do so to the ground state. Construction of antibody catalysts rests on the assumption that the difference between antibodies and enzymes is not fundamental, but merely one of specificity — over 20 years ago Jencks suggested that if antibodies could be persuaded to bind to transition states, then they should, in theory, acquire catalytic properties⁵. For the first generation of catalytic antibodies, transition-state analogues have been used for immunization in an attempt to fool the immune system into producing antibodies that possess transition-state specificity. This is followed by screening for antibodies that are catalytically active, but the process is inefficient and time consuming. One of the principal aims of current research is to reduce the amount of screening, or eliminate it completely.

Since its inception, the field has been a melting pot for exponents of immunology, chemistry and genetics — all of these techniques are being used in the drive for increased efficacy. Physical organic chemistry has so far been at the heart of things, and the synthesis of transition-state analogues remains a primary requirement. But when looking at how enzymes achieve catalysis, it is clear that transition-state stabilization alone is not sufficient, and S. J. Benkovic (Pennsylvania State University), K. Janda and R. A. Lerner (Scripps Clinic, La Jolla) presented work showing that a catalytic antibody which promotes anilide and ester hydrolysis does so by way of a putative substrate-antibody intermediate. Enzymes usually contain specific functional groups that take part in the reaction, for example carboxylate groups that act as general-base catalysts, or non-protein cofactors. Clever synthesis can select for catalytic groups in the binding

site, using antigens capable of making complementary electrostatic interactions. For instance, a positive charge introduced into the antigen will tend to be bound by antibodies that have a complementary carboxylate in their active site (K. Shokat and P. G. Schultz, University of California, Berkeley), and antibodies that catalyse β -elimination reactions have



Design of a cofactor-binding site in antibodies^{7,8}. The zinc-ligating histidine residues (silver spheres) of carbonic anhydrase were incorporated into the light chain of an anti-fluorescein antibody (grey solvent-accessible surface). The genetically remodelled antibody binds metal and antigen simultaneously, with the zinc atom (blue) positioned to interact with the fluorescein antigen (green spheres and bonds). (Computer graphics by Michael Pique, Victoria Roberts and John Tainer, Scripps Clinic.)

been generated by this process⁶. More direct ways of introducing catalytic groups are also being tried. The most common cofactor in enzyme-catalysed reactions is a metal ion, and in recent experiments^{7,8} a metal coordination site has been introduced into the light chain of an antibody, next to the antigen-binding site (Lerner and Benkovic; see figure). The metal coordination site is based on the Zn²⁺ site of the enzyme carbonic anhydrase, and comprises three histidine side chains which are incorporated into the antibody by site-directed mutagenesis. The vacant fourth coordination site will, it is hoped, be in a suitable position to interact with a bound antigen. The engineered antibody specifically binds zinc or copper ions, with an affinity constant of 10⁶ M⁻¹, and this may well be a basis for generating catalytic antibodies by combining the engineered light chain with a heavy chain that confers specificity for a given substrate. Metal cofactors have also been tried in conjunction with substrates⁹, but building them into the antibody should be of more general application. Protein engineering is also being used to improve existing catalytic antibodies (A. Plück-

thun, Max-Planck-Institute, University of Munich; P.G. Schultz).

In vivo, enzymes acquire their efficiency as a result of evolutionary processes. D. Hilvert (Scripps Clinic) has introduced the gene for a catalytic antibody that has chorismate mutase activity¹⁰ into a strain of yeast that has a defective chorismate mutase enzyme. The original antibody is a relatively poor catalyst, but by applying selection based on genetic complementation it may be possible to 'evolve' the antibody. This will provide valuable information on the evolutionary process itself, and should be a way of making sluggish antibodies more efficient.

Molecular genetics can also offer more efficient screening for catalytic activity^{11,12} (W. Huse, Ixsys Inc., San Diego; Lerner and Benkovic). Current techniques are laborious and probably sample only a small portion of the immunological repertoire, so attempts are being made to express the whole antibody repertoire in *Escherichia coli*. It should then be possible to screen thousands of antibody-producing clones with corresponding greater chances of success. Such methods hold great promise, not only for generating catalysts but for producing antibodies in general. However, they may displace the requirement for synthetic chemistry, as screening does not require a

transition-state analogue. The coming few years will show whether it proves more effective to produce antibody catalysts from naive immunological repertoires, or to obtain them using mechanism-based antigens. But in either case it is likely that further chemical or genetic modifications will be used to optimize the catalytic activity. □

Robin J. Leatherbarrow is in the Department of Chemistry, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK.

*Ciba Foundation Symposium 159 on Catalytic Antibodies, London, 1–3 October 1990. The proceedings will be published by Wiley in June 1991.

1. Tramantano, A., Janda, K. & Lerner, R.A. *Science* **234**, 1566–1570 (1986).
2. Pollack, S., Jacobs, J. & Schultz, P.G. *Science* **234**, 1570–1573 (1986).
3. Napper, A.D., Benkovic, S.J., Tramantano, A. & Lerner, R.A. *Science* **237**, 1041–1043 (1987).
4. Blackburn, M., Kang, A.S., Kingsbury, G.A. & Burton, D.R. *Biochem. J.* **262**, 381–390 (1989).
5. Jencks, W.P. *Catalysis in Chemistry and Enzymology* (McGraw-Hill, New York, 1969).
6. Shokat, K.M., Leumann, C.J., Sugawara, R. & Schultz, P.G. *Nature* **338**, 269–271 (1989).
7. Roberts, V.A. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6654–6658 (1990).
8. Iverson, B.L. et al. *Science* **249**, 659–662 (1990).
9. Iverson, B.L. & Lerner, R.A. *Science* **243**, 1104 (1989).
10. Hilvert, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4953–4955 (1988).
11. Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. & Winter, G. *Nature* **341**, 544–546 (1989).
12. Huse, W.D. et al. *Science* **246**, 1275–1281 (1989).

Giant rock avalanches

H. J. Melosh

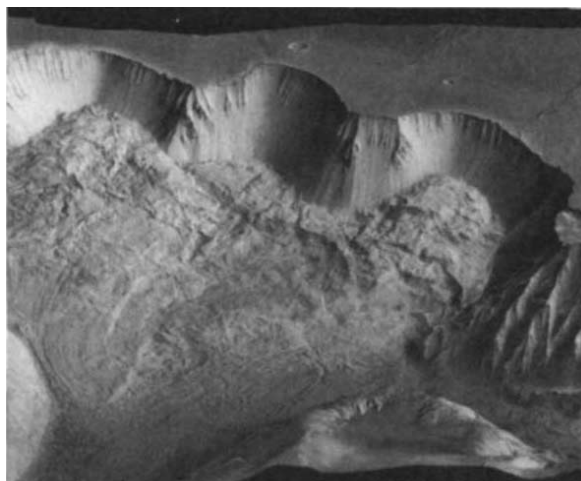
WHAT has more than a million cubic metres of rock debris, travels up to 100 km horizontally at speeds faster than 50 m s^{-1} and can climb hills? The answer to this geological riddle is a giant rock avalanche, dubbed a *sturzstrom* by the Swiss geologist Albert Heim in 1932. Known to geologists since Buss and Heim's classic study of the catastrophic rockfall at Elm, Switzerland in 1881, large rock avalanches continue to defy simple mechanical explanation. Their most surprising aspect is their high mobility: in the largest events (with volumes of 10^8 m^3 or more), dry rock debris may travel a horizontal distance 30 times farther than its net vertical fall, implying that the coefficients of friction are lower than have ever been observed for rock debris in the laboratory. Although a few *sturzstroms* have been seen fortuitously by eyewitnesses, Western geologists have mostly had to content themselves with studying their deposits. So it was a great surprise to learn that over the past few decades some of our Soviet colleagues have been actually creating *sturzstroms* for geotechnical purposes, and that in 1997 they plan to trigger a $3 \times 10^8 \text{ m}^3$ landslide in the Kirgiz Republic.

These facts, among many others, were disclosed during the course of a remarkable pair of informal workshops* attended by a small group of geoscientists from the United States, Canada, Japan and the Soviet Union, whose backgrounds ranged from field geology, through soil mechanics and the mechanics of granular flow, to explosion physics. Although the high mobility of large rock avalanches has been recognized for some time, the workshops were catalysed by new results from rather different disciplines.

Strong constraints on the mechanism of *sturzstroms* may be derived from the recent analyses of Viking images of very large *sturzstrom* deposits on Mars (see figure). Alfred McEwen (US Geological Survey, Flagstaff) provided some striking images and detailed measurements of several martian examples and showed that, for a given volume, the martian *sturzstroms* are only half as mobile as those on Earth. Nevertheless, the fact that they occur at all on Mars argues that lubrication by air or other entrained gases does not play a fundamental role in their movement. Putative *sturzstroms* on the Moon may entirely eliminate a major role for either gases or liquids. The planetary perspective was also exploited by A. V. Potapov and B. A. Ivanov (Schmidt Institute, Moscow) who showed that the collapse of

impact craters can be linked to the mechanics of large landslides, and that both phenomena have the same high mobility.

The Shasta terrane in northern California, recognized as a $4.5 \times 10^{10} \text{ m}^3$ landslide deposit in 1984, seemed to be one of the largest *sturzstrom* deposits on Earth, but J. G. Moore (US Geological Survey Menlo Park) presented evidence of sev-



A large *sturzstrom* deposit on Mars (courtesy A. McEwen).

eral 200-km-long submarine landslide deposits on the Hawaiian ridge with volumes up to $5 \times 10^{12} \text{ m}^3$ that seem to have possessed the same high mobility, as a function of size, as subaerial *sturzstroms*.

A real difficulty in studying *sturzstroms* geologists used to maintain, is that there are virtually no good exposures of the basal portions of their deposits. This commonplace wisdom was overturned by the presentation of cross-sections of the basal portions of nearly a dozen well-exposed *sturzstrom* deposits from Arizona and California (J. C. Yarnold, University of Arizona, Tucson). Such deposits turn out to be relatively common in the sediments filling enclosed basins surrounded by high mountains, but had not previously been recognized for what they are. They show intense shear and brecciation near the base of the sliding rock debris and indicate that some shear deformation occurred up to one-half to one-third of the slide's thickness from the base.

One of the great revelations of the workshops was the Soviet data base of natural, volcanic, seismogenic and artificial landslides that range in volume from 10^3 m^3 to more than 10^{10} m^3 (V. V. Adushkin, Schmidt Institute). Several *sturzstroms* with volumes of 10^6 to $8 \times 10^7 \text{ m}^3$ in Novaya Zemlya have been produced by 'point explosions' — a euphemism for underground nuclear blasts. Good before-and-after topographic data are available for these events in addition to

movies of their motion (only a sequence of stills from the movie was presented). From these data, the position of the front of the avalanche as a function of time was derived. One rather startling sequence showed material that originated at the top of the slope apparently overriding the front of the moving landslide mass.

Another surprise was the Soviet landslide dam-building programme (V. F. Korchevsky, Tashkent Institute of Hydropower and Irrigation, Uzbek Republic). For several decades, the Soviets have been creating dams with volumes of 10^4 – 10^8 m^3 by the detonation of high explosives beneath adjacent slopes. They are well aware of the high mobility of large rock masses and work this into their planning for the largest events. Preparations are now underway for a 275-m-high, 1-km-long dam with a volume of $3 \times 10^8 \text{ m}^3$ to be produced in a narrow granite gorge of the Tien Shan mountains, Kirgiz Republic, in 1997 (the Kambarata I project). For this project 250,000 tons of high explosive (ammonium nitrate and fuel oil) will be detonated in a dozen horizontal galleries driven beneath the adjacent

slopes. Studies for this project include theoretical analyses of landslide initiation, laboratory simulations using exploding air bladders in glycerol-soaked sand, and at least one small ($3 \times 10^6 \text{ m}^3$) pilot dam using 2,000 tons of explosive that was completed in 11 June 1989 at a location about 30 km west of the Kambarata I dam site. The project leaders have invited interested Western scientists to participate in studying the landslide and to make any measurements that may elucidate the mobility mechanism.

Finally, several presentations were concerned with the mechanism that imparts the observed high mobility to large masses of rock debris. My 'acoustic fluidization' mechanism can be applied to landslides, giving at least qualitative agreement with the observations although a number of crucial parameters must be determined experimentally. Computer models of high speed grain flow (C. S. Campbell, University of Southern California, Los Angeles) show very large shear and a low coefficient of friction across a narrow basal zone. Scale models using low-viscosity gelatin have been made to simulate several large natural landslides as well as the collapse of a mine dump on the Kola peninsula (S. S. Grigoryan, Moscow State University). S. Savage (McGill University, Montreal) presented a model for excess landslide run-out that involves preferential shedding of low-velocity debris and consequent maintenance of high velocity in the

*Workshops on Giant Landslides, California Institute of Technology, Pasadena, 6–7 March 1990 and Schmidt Institute of the Physics of the Earth, Moscow, 15–17 October 1990.

remaining material. One intriguing proposal (R. F. Scott, California Institute of Technology) was that a 10-m diameter, 1,000 g geotechnical centrifuge would permit model experiments on dynamically scaled analogues of natural large landslides.

In spite of the flood of new data on giant landslides, the fundamental cause of their mobility is still far from understood. It is hoped that the new opportunities for the study of their deposits both on other

planets and under the sea, and especially of the artificially triggered Soviet geotechnical landslides, will shed new light on this vexing question. It is clear that Western geoscientists have much to gain from cooperation with our Soviet counterparts in the new spirit of glasnost. □

H. J. Melosh is in the Lunar and Planetary Laboratory and Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA.

CELL BIOLOGY

Many routes lead to the pole

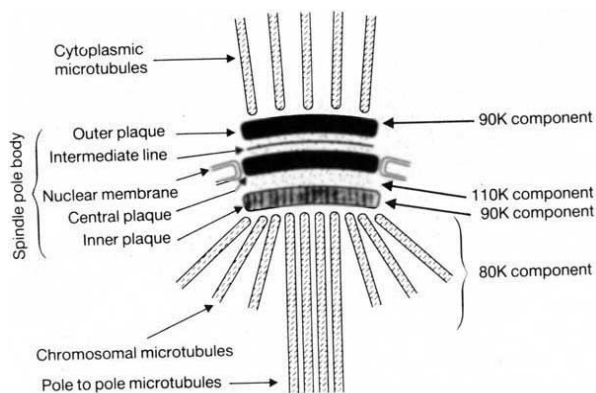
Caroline E. Alfa and Jeremy S. Hyams

MICROTUBULES in cells do not arise at random. Rather, their organization into arrays such as the mitotic spindle is choreographed by structures known as microtubule organizing centres, which were first recognized over 20 years ago¹, and which influence the dynamic properties and structural integrity of the polymers assembled onto them. The purification of organizing centres in biochemically useful amounts has proved to be incredibly difficult because they are small and generally there is only one in each cell. But now, writing in the latest issue of *Journal of Cell Biology*², Rout and Kilmartin report the isolation and characterization of a particularly well-studied microtubule organizing centre. Diagram of the localizations of the spindle pole body and the spindle pole body of the budding yeast *Saccharomyces cerevisiae*.

The spindle pole body of *S. cerevisiae* is a three-layered disk embedded in the nuclear envelope. During interphase, microtubules extend from its outer, cytoplasmic face into the growing bud. As cells enter mitosis, the pole body replicates and about 20 microtubules polymerize onto the inner, intranuclear surface of each daughter, the two sets interdigitating to form the mitotic spindle^{3,4}. The yeast spindle pole body is an attractive model for the function of microtubule organizing centres because mutants affecting its structure, replication and copy number have been isolated⁵⁻⁷. Moreover, crudely isolated pole bodies retain their capacity to nucleate microtubule assembly *in vitro*^{8,9}. But nothing is known about their biochemistry, nor of the cell-cycle regulation of their microtubule-nucleating functions.

Rout and Kilmartin released spindle pole bodies from highly purified nuclei using a combination of detergent and DNase I in low salt conditions. The resulting preparations, which retain their *in vivo*

morphology and microtubule-nucleating characteristics, were used to raise a battery of monoclonal antibodies which recognized three components of relative molecular mass 110,000, 90,000 and 80,000 (*M* 110K, 90K and 80K). The association of these proteins with the pole



body, initially confirmed by immunofluorescence light microscopy, was more accurately defined by immunogold electron microscopy. The 110K species resides between the pole body's central and innermost layers; because of its distance from the microtubule ends, it is likely to be a structural component rather than being involved in microtubule nucleation. On the other hand, the 90K species localizes at the pole body's two faces (from which microtubules grow) and is therefore a potential regulator of microtubule assembly (see figure). The 80K protein has a somewhat different distribution, and appears to be associated with the regions of spindle microtubules close to the poles rather than with the pole body itself.

Although these findings do not directly define a role for the proteins concerned, antibodies against them provide a route to the cloning and manipulation of the corresponding genes and a description of their function *in vivo* is doubtless not too far away. This is particularly timely given a number of other recent insights into the spindle pole body from less direct routes.

One of the most surprising has emerged from the study of genes concerned with another process associated with the pole body, namely, the movement of nuclei towards one another during karyogamy. One of these, *KAR1*, encodes a pole-body component of unknown function¹⁰ but another, *KAR3*, has been shown to encode a homologue of the microtubule motor protein kinesin¹¹. This colocalizes with the *KAR1* product at the outer surface of the pole body during conjugation but resides at the inner face during mitosis (M. Rose, personal communication).

In the filamentous fungus *Aspergillus nidulans* and the fission yeast *Schizosaccharomyces pombe*, genes required for the proper duplication of the pole body at mitosis also encode homologues of kinesin^{12,13}. Unlike the pole bodies in *S. cerevisiae*, those in *A. nidulans* and *S. pombe* probably act as microtubule organizing centres only in mitosis — the timing of their activation coincides with the phosphorylation of a subset of pole body proteins¹⁴ and the appearance at the pole body of the two elements of the mitotic kinase, *cdc2* and its regulatory cyclin subunit¹⁵. How (or even if) these modifications relate to the ability of the pole body to assemble microtubules is not known. But it is tempting to speculate that the link is provided by γ -tubulin¹⁶, which was first found at the pole bodies of *A. nidulans* but which is now thought to be a unique but ubiquitous component of microtubule organizing centres^{17,18}. The availability of methods to purify spindle pole bodies and of specific probes to a number of their component proteins will doubtless help investigations into these aspects of pole-body function. Equally clearly, routes to the function of microtubule organizing centres from several directions will continue to converge at the poles. □

Caroline E. Alfa and Jeremy S. Hyams are in the Department of Biology, University College London, Gower Street, London WC1E 6BT, UK.

- Pickett-Heaps, J.D. *Cytobios* **3**, 257–280 (1969).
- Rout, M.P. & Kilmartin, J.V. *J. Cell Biol.* **111**, 1913–1927 (1990).
- Byers, B. & Goetsch, L. *J. Bact.* **124**, 511–523 (1975).
- Peterson, J.B. & Ris, H.J. *Cell Sci.* **22**, 219–242 (1976).
- Baum, P., Furlong, C. & Byers, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5512–5518 (1986).
- Baum, P., Yip, C., Goetsch, L. & Byers, B. *Molec. cell Biol.* **8**, 5386–5397 (1988).
- Uzawa, S. et al. *Cell* **62**, 913–925 (1990).
- Hyams, J.S. & Borisov, G.G. *J. Cell Biol.* **78**, 401–414 (1978).
- Byers, B., Shriver, K. & Goetsch, L. *J. Cell Sci.* **30**, 331–352 (1978).
- Rose, M.D. & Fink, G.R. *Cell* **48**, 1047–1060 (1987).
- Melun, P.B. & Rose, M.D. *Cell* **60**, 1029–1041 (1990).
- Enos, A.-P. & Morris, N.R. *Cell* **60**, 1019–1027 (1990).
- Hagan, I. & Yanagida, M. *Nature* **347**, 563–566 (1990).
- Engle, D.B., Doonan, J.H. & Morris, N.R. *Cell Motil. Cytoskel.* **10**, 432–437 (1988).
- Alfa, C.E., Ducommun, B., Beach, D. & Hyams, J.S. *Nature* **347**, 680–682 (1990).
- Oakley, C.E. & Oakley, B.R. *Nature* **338**, 662–664 (1989).
- Oakley, B.R., Oakley, C.E., Yoon, Y. & Jung, M. *Cell* **61**, 1289–1301 (1990).
- Tanaka, H. et al. *J. Cell Biol.* **111**, 412a (1990).

RNA binding: β s and basics

Phillip D. Zamore, Maria L. Zapp and Michael R. Green

CONTROL of gene expression requires the interaction of regulatory proteins with specific DNA or RNA sequences. Molecular genetic and biophysical experiments have together produced a detailed picture of how proteins interact with DNA, but our understanding of how sequence-specific RNA-binding proteins, particularly eukaryotic RNA-binding proteins, recognize their target sites is rather primitive. Such recognition may involve new and different principles because RNA has the capacity for extraordinary structural diversity. A substantial advance in understanding how proteins recognize RNA is provided by Nagai *et al.* (page 515 of this issue¹), who have solved the crystal structure of the RNA-binding domain of the A protein of the U1 small nuclear ribonucleoprotein particle (snRNP).

RNA-recognition motif

The A protein contains an RNA-binding motif termed the RNA-recognition motif (RRM; for reviews see refs 2–4). The RRM consists of about 80 amino acids and contains two elements, a highly conserved ribonucleoprotein consensus sequence of eight amino acids (RNP-1), and a less well conserved sequence of six amino acids (RNP-2). Deletion analysis indicates that the entire RRM is necessary for sequence-specific RNA binding (see, for example, ref. 5). The RRM was initially recognized in poly(A) binding protein, which contains four copies of this structural motif⁶. Subsequently, the RRM was found in several prokaryotic and eukaryotic proteins that bind single-stranded nucleic acids, including heterogeneous nuclear RNP proteins, the translation factor eIF-4B, Rho protein and T4 gp32 protein of *Escherichia coli*, products of *Drosophila* genes involved in sex determination (*Sxl* and *Tra*) and cell fate determination (*elav*), three tobacco chloroplast gene products, and protein components of U1 and U2 snRNPs.

Although the RRM is recognized as the hallmark of an RNA-binding protein, it has been difficult to see how this motif provides RNA-binding specificity. Of the RRM proteins, the A protein specific to U1 snRNP has been one of the most vigorously studied. Small nuclear ribonucleoprotein particles consist of a small nuclear RNA (snRNA) tightly complexed with a set of snRNP structural polypeptides⁷. Some snRNP structural polypeptides, such as the U1 snRNP A protein, come into direct contact with specific stem-loops in the snRNA. During splicing, U1 snRNP itself binds to the 5' splice site of the primary transcript.

The crystal structure of Nagai *et al.*

reveals that the U1 snRNP A protein RRM consists of a four-stranded anti-parallel β sheet with two α helices packed against one face of the sheet. The highly conserved RNP-1 and RNP-2 motifs lie on the two central strands, at the centre of the β sheet, and probably provide general RNA-binding activity. Two aromatic residues (one in each conserved motif) project upwards from the centre of the sheet and are likely to intercalate between the RNA bases. These aromatic amino acids make contact with the RNA: the equivalent residues in another RRM protein can be crosslinked to oligodeoxythymidine⁸. A structure quite similar to that of Nagai *et al.* has been proposed from NMR spectroscopy of the U1 snRNP A protein RRM domain (J. Keene, personal communication).

Although the crystal and NMR structures show the disposition of the highly conserved RNP elements, in themselves they do not explain how RRM proteins recognize specific RNA sequences. Insight into this issue comes from experiments carried out by Mattaj and co-workers, who have delineated a protein region involved in sequence-specific binding: substitution of eight amino acids of the U1 A protein with those of the highly homologous U2 snRNP B' protein results in exchange of their RNA-binding specificities⁹.

Nagai *et al.*, resorting to molecular genetic experiments, show that of these eight amino acids, two serine residues (Ser 46 and Ser 48) are most likely to discriminate between the different binding sites of the A and B' proteins. These two serines lie on a loop containing the positively charged amino acids arginine and lysine. The crystal structure reveals a second positively charged loop positioned below and parallel to the first loop. The arginines and lysines on the two loops are essential for the protein to bind RNA. The authors speculate that the two basic protein loops 'grip' the single-stranded region of the RNA stem-loop, enabling the two serines to make the critical, discriminatory contacts.

The requirement for positively charged amino acids in binding of the U1 snRNP A protein comes as no surprise. In several well-studied DNA-binding proteins, positively charged amino acids have been shown to make contact with the DNA backbone or bases¹⁰. In fact, a separate class of RNA-binding proteins has been identified by virtue of an arginine-rich motif¹¹. These sequence-specific RNA-binding proteins, which lack the RRM, include the bacteriophage λ N protein, Tat and Rev of the human immuno-

deficiency virus type 1 (HIV-1), ribosomal and snRNP core proteins, and RNA virus coat proteins. The arginine-rich motif contains a core of six to eight basic amino acids, typically arginines, usually flanked by non-basic polar and charged amino acids. Mutational analysis shows that these arginines are necessary for RNA binding.

Arginine-rich region

Recent studies of the HIV-1 Tat protein have led to the surprising finding that the arginine-rich region alone may be sufficient for sequence-specific RNA recognition. Most strikingly, synthetic peptides consisting of 24 and 14 amino acids^{12,13} bearing the Tat arginine-rich region bind specifically to TAR RNA, the Tat recognition site.

This observation is perplexing in the light of several previous studies. For example, when the arginine-rich region of Tat, which recognizes TAR, is replaced by the arginine-rich region of the HIV-1 Rev protein, which recognizes an unrelated RNA sequence, the Tat-Rev hybrid has normal Tat activity¹⁴. This result strongly suggests that important determinants of RNA-binding specificity lie outside the arginine-rich region. Another arginine-rich protein, the bacteriophage λ N gene product, also requires the arginine-rich motif to bind RNA. But again additional sequences are required for binding specificity. Perhaps the Tat peptides do not bind with the same degree of specificity as does the intact Tat protein.

The structural versatility of RNA vastly exceeds that of DNA. The determinants of RNA-binding specificity are therefore likely to include the three-dimensional RNA structure as well as the primary sequence. Apparently the RRM motif can recognize a wide variety of RNA structures. For example, whereas the U1

1. Nagai, K., Oubridge, C., Jessen, T., Li, J. & Evans, P.R. *Nature* **348**, 515–520 (1990).
2. Mattaj, J.W. *Cell* **57**, 1–3 (1989).
3. Bradziulis, R.J., Swanson, M.S. & Dreyfuss, G. *Genes Dev.* **3**, 431–437 (1989).
4. Dreyfuss, G., Swanson, M.S. & Pinol-Roma, S. *Trends biochem. Sci.* **13**, 86–91 (1988).
5. Query, C.C., Bentley, R.C. & Keene, J.D. *Cell* **57**, 89–101 (1989).
6. Sachs, A.B., Bond, W.M. & Kornberg, R.D. *Cell* **45**, 827–835 (1986).
7. Steitz, J.A. *et al.* in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (ed. Birnstiel, M.L.) 115–154 (Springer, Heidelberg, 1988).
8. Merrill, B.M., Stone, K.L., Cobiach, F., Wilson, S.I.I. & Williams, K.R. *J. biol. Chem.* **263**, 3307–3313 (1988).
9. Scherly, D., Boelens, W., Dathan, N.A., van Venrooij, W.J. & Mattaj, J.W. *Nature* **345**, 502–506 (1990).
10. Harrison, S.C. & Aggarwal, A.K. *Rev. Biochem.* **59**, 933–969 (1990).
11. Lazinski, D., Grazdzielska, E. & Das, A. *Cell* **59**, 207–218 (1989).
12. Roy, S., Delling, U., Chen, C.-H., Rosen, C.A. & Sonenberg, N. *Genes Dev.* **4**, 1365–1372 (1990).
13. Weeks, K.M., Ampe, C., Schultz, C.S., Steitz, T. & Crothers, D. *Science* **249**, 1281–1284 (1990).
14. Subramanian, T., Kuppuswamy, M., Venkatesh, L., Srinivasan, A. & Chinnadurai, G. *Virology* **176**, 178–183 (1990).
15. Rould, M.A., Perona, J.J., Söll, D. & Steitz, T.A. *Science* **246**, 1135–1142 (1989).

snRNP A protein binds to an RNA stem-loop, the RRM-containing poly(A) binding protein recognizes a single-stranded stretch of adenosines. To 'read' the primary RNA sequence, RRM proteins may distort the structure of their RNA ligands. Such distortions occur, for example, in transfer RNA^{Gln} following binding of glutamyl-tRNA synthetase¹⁵.

Ultimately, the relative contributions of primary sequence and structure in the recognition of RNA by proteins will be

resolved by solving co-crystal structures. But until then we expect that molecular genetic approaches, using altered RNAs and proteins, will continue to provide important clues about the rules governing RNA-protein interactions. □

Phillip D. Zamore, Maria L. Zapp and Michael R. Green are in the Program in Molecular Medicine, University of Massachusetts Medical Center, 373 Plantation Avenue, Worcester, Massachusetts 01605, USA.

COMETARY ORIGINS

From dust to dust

Clark R. Chapman

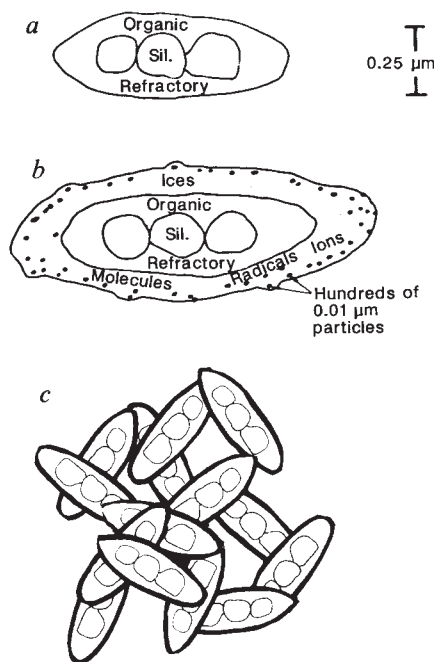
FREEZE a collection of interstellar dust into a porous, icy ball, then hurl that ball close to the Sun so that the ices evaporate, liberating the dust. Do the spectra of such interplanetary dust particles streaming away from the comet resemble those of interstellar dust — the hypothesized primordial components of comets? That is the question J. Mayo Greenberg and J. I. Hage ask in new work appearing in the *Astrophysical Journal* (361, 260–274; 1990) on the origin of comets.

Little is known about comets, even following the detailed investigations of Halley a few years ago by several spacecraft. But comets are widely regarded as being the most undisturbed, primitive remnants of the swarms of planetesimals from which the outer planets of our Solar System accreted. Of course, comets are not unmodified. Most well-observed comets, including Halley's, have been around the Sun numerous times, resulting in modification of their surface layers, at the very least.

The holy grail in cometary science is the notion that comets consist of relatively unaltered interstellar grains. The Solar System is believed to have been born through the collapse of a cloud of such grains and the gas associated with them, and comets may have formed directly from gravitational instabilities in the outer reaches of the solar-nebular disk. Alternative models envisage that heating processes in the solar nebula would have altered or even evaporated the grains before they accreted to form larger bodies, so that comets would be made of nebular condensates, rather than from the original interstellar grains. Individual comets may have undergone still further modification, akin to the thermal metamorphism and collisional evolution that has affected many asteroids.

Mayo Greenberg, of the University of Leiden, has long investigated the hypothesis that comets are aggregates of interstellar grains. He was prominent in demonstrating that such grains would consist, originally, of silicate cores surrounded

by icy mantles. Subsequent photolysis by irradiation (by ultraviolet light) in the interstellar environment would have converted the icy mantles into more refractory compounds rich in organics. In the developing solar nebula, the core-mantle



Greenberg's picture of (a) a grain of interstellar dust, (b) a precometary grain and (c) a coma grain with a porosity of 0.93–0.98. (From Greenberg & Hage *Astrophys. J.* 361, 260–274; 1990.)

grains themselves would become coated again by ices.

Over the years, Greenberg has elaborated on his notion of comets as low-density aggregates of such particles, as shown in the figure. Such comets would be very porous, resulting in very low densities for cometary nuclei. The results from the Halley encounter seemed to confirm elements of Greenberg's hypothesis. No less an authority than the creator of the dirty-snowball model for cometary nuclei, Fred L. Whipple, concluded that "the general nature of the particles is consistent

with Greenberg's (1982) theory that they are indeed made by interstellar dust and that '~20% of comets consists of complex non-volatile organic molecules of prebiotic type'" (*Astr. Astrophys.* 187, 852–858; 1987). Whipple refers to the so-called CHON particles (named after the common organic elements) that were a significant finding of the Halley missions.

Neither the distribution of particle sizes in Halley's coma, nor the bulk density of its nucleus has succumbed to analysis, however. These are the issues tackled by Greenberg and Hage in their new paper. They have developed a model, including the composition, size and packing densities of the particles, for understanding ground-based observations of Halley as well as those gathered during the encounter with observations of interstellar dust. They use it to calculate the temperatures and emission characteristics of these amorphous dust aggregates. The observations to be matched include the shapes and strengths of two infrared emission bands in the comet's spectrum.

In short, Greenberg and Hage conclude that the interstellar dust model of comets is consistent with their quantitative model of Halley's spectrum and with other known facts (at least ones that they choose to believe) about Halley and cometary meteors. The question is whether or not astronomers should be similarly impressed. It is unlikely that many more definitive data will be available for some time, as even those comet missions now under study and development (such as the Comet Rendezvous Asteroid Flyby, or CRAF, mission) will not study a comet for many years.

There are reasons to be a bit sceptical. There have been continuing uncertainties about the actual size distribution of Halley dust. In any case, such dust is not precisely what Greenberg and Hage have modelled: their comet consists of a spherical aggregate of identical particles, each with a silicate core of radius 0.075 micrometres. (They consider an additional population of much tinier carbonaceous particles with radii below 0.01 micrometres, probably incorporated into the icy mantles of the larger precometary grains.)

Astronomical remote-sensing measurements of size distributions in the past have been fraught with uncertainties. The particles in Saturn's rings were commonly modelled as having a modal size distribution, and were fitted to telescopic data at visible wavelengths. As the rings became probed by longer and longer wavelengths (infrared, radar, and finally the radio sounding by Voyager), it became clear that a quasi-power-law size distribution was involved with the dominant mass residing in decametre (house-sized) objects.

Our understanding of the interplanetary particle population is undergoing simi-

lar revisions as we probe it at different wavelengths. Although Greenberg relies (probably correctly) on the inferred properties of 'cometary' meteors, there is an increasing fraction of interplanetary particles that are now thought of as being of asteroidal origin. Comet trails, observed by the Infrared Astronomical Satellite (IRAS) and by ground based radar, are turning up ever larger particles. Interstellar particles, of course, cannot be probed by radar. But I wonder if the simple models that astronomers and astrophysicists have laboured to infer from their data might suffer from an analogous blindness owing to the inevitable limitations in observing these distant environments.

Halley's Comet, of course, is just a single object, and it has no doubt evolved from its original condition since it first

entered the inner Solar System. There are contradictory indications about whether most comets are essentially identical in physical and chemical constitution or whether they are as varied as the asteroids are. It will behove us to study more comets — both with telescopes and by spacecraft visits — before we ascribe to all comets the characteristics of Halley's. Meanwhile, Greenberg's elaborate model for comets as aggregates of interstellar grains remains one of the paradigms in this field. Despite my nitpicking, Greenberg and Hage's work is the most rigorous demonstration to date of the association between inferred properties of interstellar grains and the dust in our best-observed comet. □

Clark R. Chapman is in the Planetary Science Institute, Tucson, Arizona 85719, USA.

FLUID DYNAMICS

Separation in flowing fluids

Daniel D. Joseph

Two liquids of different viscosities will stratify with the heavy liquid below, when stationary. But when these stratified liquids are made to flow down a pipe, the less viscous liquid will tend to encapsulate the more viscous liquid, even lubricating it, regardless of their relative densities (Fig. 1). Calculations on page 523 of this issue try to make sense of this behaviour.

The segregation of two liquids by viscosity which manifests itself in pipes as encapsulation is one example of the tendency of the low-viscosity constituent to migrate into regions of high shear. There are many applications, current and potential, for this trait, ranging from the coextrusion of molten plastics, to segregation in forming layered materials of superior strength, to the water-lubricated

transport of viscous liquids and dense suspensions of solid particulates. Water-lubricated pipelining of crude oils is one emerging technique by which drag on the oil flow can be reduced by factors of the order of the viscosity ratio, 10^{-4} or less (C. L. Merkle & S. Deutsch in *Viscous Drag Reduction in Boundary Layers* 123, (eds D. M. Bushnell & J. N. Hefner) 396–402 (*Am. Inst. Aeronaut. Astronaut.*, 1990)).

Many different flow configurations occur in such flows, each giving rise to a different drag. It is possible to see perfect core-annular flows, bamboo and corkscrew waves of oil in water (Fig. 2), or slugs, bubbles and dispersions of water in oil. Understanding how one flow evolves to another is a mystery of fluid dynamics whose solution would help in the optimization of lubrication. The evolution can be formulated as an initial-boundary-value problem for two immiscible liquids satisfying the Navier-Stokes equation and interface conditions between liquids. This difficult study is at best computationally intensive and at worst either too expensive or even beyond the capabilities of modern supercomputers.

The mechanisms underway in the flows which segregate by viscosity can be partly understood by studies of the stability of core-annular flows and more directly by variational principles. For example, it has been postulated that the realized configuration of flows of two liquids is the one that minimizes the viscous dissipation of energy. In fact, several different versions of this principle put the more viscous constituent in the core.

H. W. Stockman, C. T. Stockman and C. R. Carrigan introduce elsewhere in this issue (*Nature* 348, 523–525; 1990) an entirely new method for studying viscous

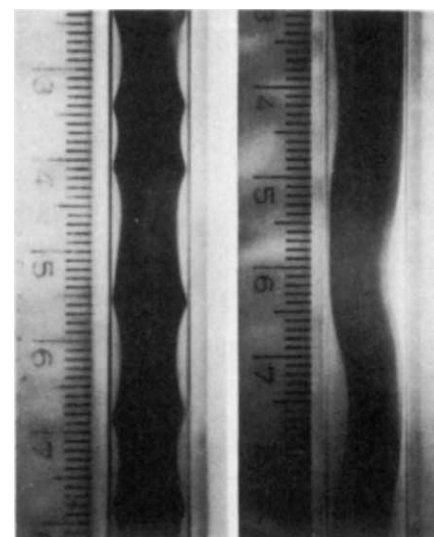


Fig. 2. Water-lubricated pipelining of a 14-poise oil in water in a vertical U loop. Bamboo waves (left) arise in the up flow of the U loop where buoyancy stretches the oil. Corkscrew waves (right) arise from the buckling under compression of gravity of water in oil in down flow, where the gravity and the pressure gradient are opposed.

segregation and the evolution of lubricating flows. They adapt a method of cellular automata to follow the motions of incompressible liquids with different viscosities and interfacial tension in two-dimensional simulations of the initial-value problem. They show how the lubricating configuration with the low-viscosity liquid on the walls can arise. Their method is not, and is not meant to be, an exact description of the flow's dynamics: it models the Navier-Stokes equation only in an average sense. It is greatly encouraging that several variants of the rules for the automata lead to identical results and that they all give rise to lubrication. We can hope that the exact mechanisms which lead to lubrication or its failure may be illustrated by studies of motions of the lattice gas, even if the methods of cellular automata for quantitative prediction are never realized.

The manner in which high-viscosity liquids are lubricated depends strongly on conditions and could never be described by one variational principle. A kind of anthropomorphic variational principle, which like other principles may be believed but not proved, goes like this: "High-viscosity liquids are lazy. Low-viscosity liquids are the victims of the laziness of high-viscosity liquids because they are easy to push around." Perhaps the methods of cellular automata introduced by Stockman *et al.* can help us to better understand and control the lazy fellow and his victim. □

Daniel D. Joseph is in the Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, Minnesota 55455, USA.

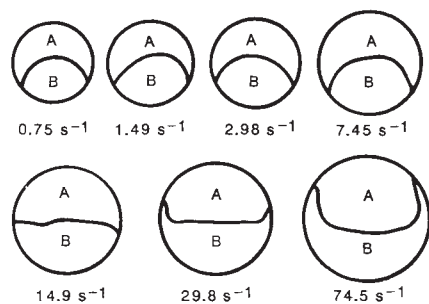


Fig. 1 Fluids A and B are viscoelastic fluids, and their viscosities vary depending on the apparent wall shear rate (given here in reciprocal seconds). In the upper set of arrangements, fluid B is the more viscous, and fluid A tends to wrap around fluid B. In the lower left, the viscosities of the fluids are about equal and the flat interface is retained. In the lower middle, fluid A is the more viscous and fluid B tends to wrap around fluid A. In the lower right arrangement, fluid A is the more viscous. (From J. H. Southern & R. L. Ballman *Appl. Polymer Symp.* 20, 175–189; 1973.)

RÉSUMÉ

Stop making sense

DURING mitosis, the chromosomes in the nucleus of a cell move towards the poles of the mitotic spindle, a scaffold of microtubules, in preparation for nuclear and cell division. R. Margolis and collaborators have identified a protein from brain that stabilizes the dynamic structure of the spindle and may generate the movement of chromosomes to the spindle's poles *in vivo* (*EMBO J.* **9**, 4095–4102; 1990). The protein, called 'stop' (stable tubule only polypeptide), is known to stabilize microtubules *in vitro*. The new work shows that it is tightly associated with microtubules *in vitro* during three cycles of assembly and disassembly. The authors also show that, under physiological conditions, stop is preferentially distributed only on those microtubules thought to move chromosomes to the poles in anaphase. The next step is to see if stop could have this function in cells other than neurons.

See-through ice

ANTARCTIC sea ice is at its most transparent in the spring (September–November), just when ozone depletion allows the highest fluxes of solar ultraviolet radiation, according to H. J. Trodahl and R. G. Buckley (*Geophys. Res. Lett.* **17**, 2177–2179; 1990). Concerned by the deleterious effect of increased ultraviolet levels on Antarctic ecology, the authors monitored the ice's transparency using an artificial light source. They estimate that the severe ozone holes of 1987 and 1989 would have caused a tenfold increase in the dose of ultraviolet to algae below the ice. The result, they say, is that the Antarctic ecosystem must have been affected, as these algae are highly sensitive to this radiation.

Flipping lipids

THE inner and outer leaflets of the plasma membranes of red cells — and many others — have different lipid compositions. The asymmetry is maintained by an ATP-dependent translocase. The hunt for this enzyme has converged recently on the Rhesus blood group protein. A.J. Schroit *et al.* (*Biochemistry* **29**, 10303–10306; 1990) have now shown that a protein, which migrates in gels with the Rhesus polypeptide, is labelled with a lipid affinity analogue or with an inhibitor of translocation, and the labelled molecule is precipitated with anti-Rhesus antibody. So far so good — but both Schroit *et al.* and R.E. Smith and D.L. Daleke (*Blood* **76**, 1021–1027; 1990) have looked at Rh_{null} cells, which are devoid of the Rhesus protein, and found that they translocate lipids just like normal cells. This leads to a fine philosophical antithesis, for Smith and Daleke conclude that the translocase cannot be the Rhesus protein and Schroit *et al.* that the Rhesus protein must be present in Rh_{null} cells in some new, antigenically silent phenotype.

SOLAR PHYSICS

The solar — stellar connection

Mark S. Giampapa

THE 11-year solar cycle, it is currently held, has just passed its maximum — a maximum more intense than seen for many decades. But how big is the range of physical behaviour that the Sun can go through? The 'Maunder minimum', a period of quiescence between 1650 and 1730, shows that activity outside the present range is possible. On page 520 of this issue¹, Baliunas and Jastrow tackle this question by looking at the range of activity apparent on other stars that are of the same type as the Sun. Given the Sun's influence on our climate, their evidence for deep Maunder-minimum-type quiescence is particularly interesting.

The Sun's atmosphere reveals a variety of bright features. They originate from magnetic fields generated deep in the solar interior. Dynamo processes are presumed to give rise to the intense chromospheric and coronal emissions which often coincide with the sites of magnetic field structures. Stars also show these characteristics, although there can be important differences in the scale of activity associated with emergent magnetic flux. Just as detailed study of the Sun serves as a guide to stellar magnetism, so the parallel investigation of stellar analogues of solar magnetic activity yields tests of theories developed purely within a solar context. Quantities such as rotation rate, depth of the convection zone, surface gravity and effective temperature and evolutionary status can take a wide range of non-solar values, stretching models well beyond any familiar territory.

Although the energy residing in solar and stellar surface magnetic fields is negligible compared to the energy associated with total luminosity (as is the radiant energy of the associated hot atmospheric layers above the photosphere that we identify as the chromosphere and corona), the dynamical role the fields play is significant. In particular, magnetic fields determine the angular-momentum evolution of the late-type stars and modulate their total radiative output. In the case of the Sun, this leads to observable changes in the total irradiance.

As demonstrated by Baliunas and Jastrow, solar-type stars can undergo a much wider range of cyclical variation than is known for the Sun. The interpretation of these chromospheric variations within the context of changes in magnetic activity relies on the assumption that chromospheric emission features are correlated with surface magnetic flux. This is well-established in the case of the Sun¹. Direct evidence has recently emerged that this is also true for solar-like stars. In particular, enhanced levels of chromo-

spheric and coronal emission in solar-type stars are correlated with high fractional coverage (or filling factor) of magnetic field¹. Thus, high intensities of chromospheric Ca II (Ca⁺) emission correspond to high levels of surface magnetic flux. The direct correlation of luminosity variability in solar-type stars with changes in chromospheric emission not only corroborates the detection of subtle variability in the solar irradiance but strongly suggests that the range of solar variability can be larger⁵. The potential range depends on the level of magnetic flux present at each phase of the solar cycle.

Another of Baliunas and Jastrow's results, indicating that solar-type stars have extended periods of dormancy in magnetic activity, is particularly exciting. The historical record⁶ suggests that there have been periods, such as the Maunder minimum, of abnormally low levels of magnetic activity on the Sun in the past, but the conclusion is controversial. When these stellar observations are combined with the extant ¹⁴C terrestrial record⁷, another indicator of solar activity, the evidence for episodes of nearly complete solar quiescence becomes quite persuasive. Moreover, Baliunas and Jastrow find that the level of chromospheric emission in quiescent stars is even lower than that observed during the minimum phase of the modern solar cycle. The results tentatively suggest that solar-like stars spend prolonged periods in these minima. In view of the correlation of brightness changes with the level of magnetic activity as seen in both solar and stellar data, the impact of a Maunder-minimum phase on the terrestrial climate could be significant.

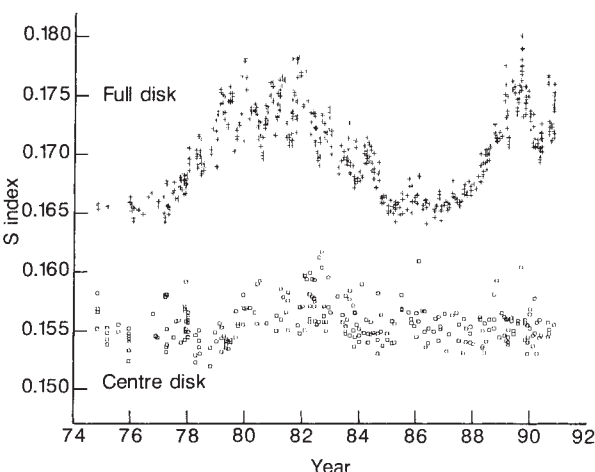
A comparison of the strength of the solar Ca II line emission strength as seen in full disk observations and in centre-of-the-disk observations (see figure) confirms that the disk-integrated Ca II K line of the Sun (the Sun seen as a star) exhibits the 11-year solar cycle, but that variations at disk centre do not. Thus, the cyclic variability in radiative proxies of magnetic activity must arise from localized concentrations of magnetic flux rather than from

1. Baliunas, J. & Jastrow, R. *Nature* **348**, 520–523 (1990).
2. Vaiana, G. S. & Rosner, R. A. *Rev. Astr. Astrophys.* **16**, 393–428 (1978).
3. Skumanich, A. *et al.* *Astrophys. J.* **200**, 747–763 (1975).
4. Saar, S. H. in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) (University of Arizona Press, Tucson, in the press).
5. Radick, R. R. in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) (University of Arizona Press, Tucson, in the press).
6. Eddy, J. A. in *The Solar Output and its Variation* (ed. White, O. R.) 51–71 (Colorado Associated University Press, Boulder, 1977).
7. Damon, P. E. in *The Solar Output and its Variation* (ed. White, O. R.) 429–448 (Colorado Associated University Press, Boulder, 1977).

subtle global effects. Furthermore, the K-index for the centre of the disk is systematically weaker than the full-disk measurement at solar minimum. Evidently, the disk-centre observations correspond closely to the non-magnetic component of the quiet Sun. Hence, the level of emission in quiet regions at disk's centre must be representative of the degree of activity (actually, 'inactivity' would be more appropriate) that would be expected during a Maunder minimum phase. In fact, an approximate conversion between the Mount Wilson 'S-index' used by Baliunas and Jastrow and the K-index measured by Livingston and White indicates that the quiescent disk-centre observations are very similar to the mean value of the S-index for the magnetically flat stars in the sample considered in the current work.

Other crucial solar parameters, such as the effective temperature, may also vary. Observations⁸ of the temperature-sensitive neutral carbon line at 538.0 nm, which is formed in the solar photosphere, show that this feature has been getting steadily stronger over the past 12 years. The increase seems to be unmodulated by the activity cycle, and indicates a temperature rise of 4.6 K. This value, in turn, implies an increment in the solar constant (about 0.3 per cent) which is greater than that seen by the ACRIM experiment (about 0.1 per cent). The photospheric model used for the carbon line may be wrong, of course, or the change in the strength of this feature may be due to other factors, such as microturbulent velocity fields, in addition to variations in effective temperature.

By contrast, other investigations⁹ based on sensitive measurements of the limb brightness suggest that a substantial fraction of the variation in the solar irradiance during the cycle can be attributed to temporal changes in the latitude-dependent surface temperature of the Sun. This approach depends upon the application of



The strength of the Ca II K line as measured by W. C. Livingston and O. R. White using the McMath telescope of the National Solar Observatory. The measurements have been approximately converted to the 'S-index' used at Mount Wilson. The full disk data denoted by (+) clearly show the solar cycle. The disk-centre observations are represented by boxes. The data obtained before 1984 are observations of the geometrical centre of the disk. Following 1984, the data are nominal disk-centre observations — that is, the 2 arcminute aperture was moved around slightly to isolate the K index for the most quiet area, uncontaminated by magnetic active regions, in the neighbourhood of disk centre.

a model-dependent correction for the contributions by spot complexes and faculae to the observed flux variability.

On stellar evolutionary timescales, we know that young stars are characterized by enhanced levels of magnetic-field-related chromospheric and coronal emission with X-ray and extreme-ultraviolet fluxes 100–1,000 times the present solar values¹⁰. These enhanced levels of activity decay with time accord-

ing to an exponential or a power-law dependence¹¹. If the Sun followed the same path of evolution, then its irradiance in the ultraviolet was once substantially greater than today. These results have been used in exploring the atmosphere of the young Earth, especially with respect to the formation of free oxygen and a protective ozone layer in advance of biological activity¹². Among the key questions that arise, however, is whether the Sun actually passed through this 'T Tauri' stage of evolution and what the state of the Earth might have been at the time. Stellar observations provide circumstantial evidence that the young Sun did experience a T Tauri phase; certainly, the very existence of our Solar System is consistent with the formation of a disk around a pre-main sequence Sun, as observed in many T Tauri stars. Irradiation records locked within meteorites contain evidence for an enhanced level of flare activity in the solar past though the question still remains open¹³.

Because ultraviolet variability contributes significantly to solar, and probably stellar, total irradiance variations¹⁴, and because energy at these wavelengths is screened from us by the ozone layer, any further elucidation of the cycles will require space-based observations. Indeed solar ultraviolet variations may have important consequences for the terrestrial climate and, if for that reason alone, the solar–stellar connection will continue to be explored. □

Mark S. Giampapa is in the National Solar Observatory, National Optical Astronomy Observatories, PO Box 26732, Tucson, Arizona 85726, USA.

ANION TRANSPORT

Revelations of a chloride channel

Harvey F. Lodish

THE cloning and sequencing of the first voltage-gated chloride channel, reported by Jentsch and colleagues on page 510 of this issue¹, is notable for several reasons. The electroplax channel of the electric ray (*Torpedo*) defines a new class of membrane channel proteins; in particular, its sequence is unrelated to that of glycine- or GABA-triggered Cl⁻ channels or to voltage-gated sodium or potassium channels. The report illustrates the power of strategies based on functional expression for cloning rare messenger RNAs whose protein products are not purified or even identified. And the channel's complementary DNA may provide an entrée into identification and cloning of other Cl⁻ channels, particularly the one(s) in airway epithelial and other cells whose regulation is altered in cystic fibrosis.

Why is the electric organ of *Torpedo* so rich in Cl⁻ channels? Jentsch *et al.* estimate that one mRNA per 1,000 in the cells codes for the cloned channel. The organs consist of stacks of modified muscle cells that have lost contractile fibres. The 'top' surface of each cell is innervated by cholinergic neurons, whose stimulation activates nicotinic acetylcholine receptors, causing Na⁺ entry and depolarization of (mainly) the 'top' plasma membrane. The 'bottom', or contra-innervated, plasma membrane, in contrast, is abundant in Cl⁻ channels. The increased permeability of this membrane to Cl⁻ balances the entry of Na⁺, and 'clamps' the potential of the 'bottom' membrane close to its resting potential of 90 mV, inside negative.

The combination of depolarization of the 'top' membrane (becoming slightly

8. Livingston, W. C. in *Climate Impact of Solar Variability* (eds Schatten, K. H. & Arking, A.) 336–340 (NASA CP 3086, Washington DC, 1990).

9. Kuhn, J. R., Libbrecht, K. G. & Dicke, R. H. *Science* **242**, 908–910 (1988).

10. Giampapa, M. S. in *Solar Radiative Output Variations* (ed. Foukal, P.) 241–249 (CRI, Cambridge, Massachusetts, 1987).

11. Walter, F. M. & Barry, D. C. in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) (University of Arizona Press, Tucson, in the press).

12. Canuto, V. M. *et al.* *Nature* **305**, 281–286 (1983).

13. Caffee, M. W. *et al.* in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) (University of Arizona Press, Tucson, in the press).

14. Lean, J. *Science* **244**, 197–200 (1989).

inside-positive) and normal polarization of the bottom (inside-negative) results in a voltage gradient across each cell, top-to-bottom, of about 90–100 mV ('bottom-positive'). The stacking of hundreds of such cells, like batteries in a series, generates the stunning potential of about a hundred volts characteristic of these organs. The abundance of acetylcholine receptors in the cells was essential to their cloning², and similarly Jentsch took advantage of the abundance of Cl^- channels in using them as the source of mRNA.

The electroplax Cl^- channel, like the erythrocyte anion exchange protein, band 3, is inhibited by the drugs based on modified stilbene disulphonates such as DIDS. However, a relatively high concentration — in the micromolar range — is required for inhibition. Surprisingly, the two main DIDS-binding proteins in electroplax membranes are the α subunit of the ($\text{Na}^+ + \text{K}^+$) ATPase and a second protein Jentsch and colleagues cloned a few years ago but which is not a subunit of a Cl^- channel³. This highlights the potential hazards in using an inhibitor of relatively low affinity and specificity to purify a channel (or any other) protein.

The cloning strategy that was ultimately successful involved functional expression of the protein in *Xenopus* oocytes. Injection of total electroplax mRNA induces expression of a voltage-sensitive Cl^- channel, and Jentsch *et al.* selected cDNAs that, when hybridized to total mRNA, removed (hybrid-depleted) the channel-inducing activity. Moreover, an RNA transcript of the full-length cDNA *in vitro* induced expression of Cl^- channel activity. The encoded protein of 805 amino acids, has 12 (or 13) presumed membrane-spanning α helices and in sequence does not resemble any known protein.

The cloned channel protein is voltage-sensitive; that is, it is slowly activated when the membrane is hyperpolarized. Perhaps surprisingly, the channel does not contain a sequence resembling the presumed 'voltage gating helix' (helix 4) in voltage-dependent K^+ and Na^+ channels, where every third residue is arginine or lysine⁴. However, these cation channels open when the membrane is depolarized.

The cloned channel protein has many properties in common with the major electroplax Cl^- channel studied in detail by Miller and his associates in reconstituted lipid bilayers⁵. In particular, once the channel is opened the flux of Cl^- ions increases with membrane depolarization. Single-channel recording will be required to determine whether the two are identical (there may be a regulatory subunit). Both the electroplax Cl^- channel and one from kidney have two Cl^- diffusion pathways, and the electroplax channel is thought to be a homodimer. The opening and inactivation of the two Cl^- 'proto-channels' are coupled to the chloride transmembrane

electrochemical gradient — a novel mechanism of channel gating⁶ — and it will be of some interest to see whether this property is intrinsic to the cloned Cl^- channel protein.

An important question is the relationship of this channel to the apical Cl^- channel in airway and sweat duct (and other) epithelial cells which is defective in cystic fibrosis (CF)^{7,8}. The cloned 'CF gene' product is almost certainly not a Cl^- channel, but a protein that regulates channel activity, even though expression of a wild-type gene in a CF epithelial cell reverses the defect in Cl^- transport^{9,10}.

Can one use molecular hybridization with the electroplax channel cDNA to clone the mammalian 'CF Cl^- channel'? Can it be used to clone other interesting Cl^- channels, such as those in coated vesicles¹¹ or in the apical membrane of the gastric oxyntic cell¹² that, together with an ATP-driven H^+ pump, are required to establish a pH gradient? Attempts of this sort with other ion transport proteins might be instructive. When Kopito and I cloned the erythrocyte anion exchange protein¹³, we thought we could use the cDNA as the probe to isolate clones for anion channel proteins. This approach did yield two other cDNAs encoding anion exchangers but no anion channels¹⁴. The cDNA encoding the *Drosophila Shaker* voltage-sensitive K^+ channel has been used to clone a large number of insect and mammalian K^+ channel proteins¹⁵. But this large 'superfamily' does not include two key types of K^+ channel proteins, those activated by ATP or by Ca^{2+} ions. Thus, once one clones a new type of channel protein or transporter, one can rather quickly clone its close relatives, yet increasingly sophisticated expression cloning strategies will be needed to clone entirely new types of membrane proteins. How many families of anion channels exist is an open and pressing question. □

Harvey F. Lodish is in the Whitehead Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA.

1. Jentsch, T.J., Steinmeyer, K. & Schwartz, G. *Nature* **348**, 510–514 (1990).
2. Noda, M. *et al.* *Nature* **302**, 528–532 (1983).
3. Jentsch, T.J., Garcia, A.M. & Lodish, H.F. *Biochem. J.* **261**, 155–166 (1989).
4. Jan, L.Y. & Jan, Y.N. *Cell* **56**, 13–25 (1989).
5. Miller, C. & Richard, E.A. in *Chloride Channels and Carriers in Nerve, Muscle, and Glial Cells* (eds Alvarez-Leefmans, F.J. & Russel, J.M.) 383–405 (Plenum, New York, 1990).
6. Richard, E.A. & Miller, C. *Science* **247**, 1208–1210 (1990).
7. Quinton, P.M. *FASEB J.* **4**, 2709–2717 (1990).
8. Welsh, M.P. *FASEB J.* **4**, 2718–2725 (1990).
9. Rich, D.P. *et al.* *Nature* **347**, 358–364 (1990).
10. Drumm, M.L. *et al.* *Cell* **62**, 1227–1233 (1990).
11. Xie, X.-S., Crider, B.P. & Stone, D.K. *J. Biol. Chem.* **264**, 18870–18873 (1989).
12. Demarest, J.R., Loo, D. & Sachs, G. *Science* **245**, 402–404 (1989).
13. Kopito, R.R. & Lodish, H.F. *Nature* **316**, 234 (1985).
14. Lindsey, A.E. *et al.* *Proc. Natn. Acad. Sci. U.S.A.* **87**, 5278–5282 (1990).
15. Jan, L.Y. & Jan, Y.N. *Trends Neurosci.* **13**, 415–419 (1990).

Only connect

BRITAIN, like many other countries, suffers chronically from road congestion. Existing roads can be upgraded and new ones built, but every stretch of extra road is jam-packed from the moment it is opened. Daedalus now proposes to double or triple our road capacity almost at a stroke.

He points out that in a fast stream of traffic, each car is separated from the next by at least ten car-lengths of empty road. This gap gives each driver the chance to see and react to the brake lights of the car ahead, should it decelerate unexpectedly. Simply by halving the reaction time of the drivers, the cars could pack twice as tightly, and road capacity would double.

The obvious technology for this job is the short-range infrared link used to control television sets. An infrared brake light, coded not merely to say that the brake was on, but to relay the changing speed and deceleration of the vehicle, could easily be detected by the car behind. Faster than any human driver, it could apply the appropriate brake force to avoid collision. In the process, of course, it would transmit its actions to the car behind it, and the message would pass down the traffic stream. A forward-shining infrared 'acceleration light', telling each car the speed and acceleration of the one behind, would complete the system. The stream of traffic would become a virtual train, its cars coupled tightly by invisible elastic couplings.

The technology should be quite simple. Many modern cars have power brakes and electronic engine management systems that lend themselves to remote control; the new fittings would be mainly cheap electronics with few expensive mechanics. Two vehicles equipped with the system would establish communication by a 'handshake' of the sort used by fax machines, and would then close up to their safe distance. A dashboard light would warn each driver of the connection, and his instinctive alarm at the tight spacing should soon wear off. Drivers should soon learn to switch rapidly between normal and 'automatic' safe distance perception, just as they switch rapidly between the different safe driving rules for single and multiple carriageways.

Once brought to market, the new system should spread fast. Drivers without it, painfully aware that their car was being shunned by those around, will be shamed by this continuous public demonstration of their disregard for safety, wasteful use of road space, and lack of the latest motoring gadget. All over our newly capacious motorway network, the 'trains' of self-satisfied, tight-packed motorists will increase in length and frequency. Soon only the odd unmodified banger will stagger along in enforced isolation, a pariah of the roads.

David Jones

491

arm loops is the same. The molecules are asymmetric because the side arm is inserted in the left-hand sequence of the stem in one and the right-hand sequence of the other. Alternatively, the side arm could be inserted at the same point but with sequences with opposite 5' → 3' and 3' → 5' polarities. In either case, the orientation of the bases in the terminal stem loop is different with regard to the orientation of those in the side arm loop. It is therefore possible to envisage a protein which might bind to these two loops of one molecule, but not those in the other. This could generate handedness or asymmetry at the molecular level.

I have suggested elsewhere³ that secondary and tertiary RNA structures, encoded by maternal DNA, may be important cytoplasmic determinants in the egg and

developing embryo. One of the best-known examples of asymmetry is the direction of coiling of the shell of snails, and in *Limnaea* this is known to be inherited maternally^{4,5}. With regard to genetic control, only one or other of asymmetric molecules like those in the figure might be transcribed from DNA, and mutations could lead to reversal or loss of asymmetry.

ROBIN HOLLIDAY

CSIRO Laboratory for Molecular Biology,
PO Box 184, North Ryde,
New South Wales 2113,
Australia

1. Galloway, J. *Nature* **346**, 223–224 (1990).
2. Brown, N.A. & Wolpert, L. *Development* **109**, 1–9 (1990).
3. Holliday, R. *New Biologist* **1**, 337–343 (1989).
4. Boycott, A.E. & Diver, C. *Proc. R. Soc. B* **95**, 207 (1923).
5. Sturtevant, A.H. *Science* **58**, 269–270 (1923).
6. Boycott, A.E., Diver, C., Garstang, S.L. & Turner, F.M. *Phil. Trans. R. Soc. B* **219**, 51–131 (1930).

Earthquake lights and seismicity

SIR—Sightings of earthquake lights and other similar anomalous phenomena in the night sky have been reported for many centuries. But it is only recently that such controversial phenomena have been linked to the geological features and tectonic processes underlying earthquake activity^{1,2}.

Earthquake light (EQL) sightings reported between 1 November 1988 and 21 January 1989 in the Saguenay region 200 km north of Quebec city, Canada (Fig. 1) were found to be associated with 54 seismic shocks³. The seismic activity was recorded by a network of portable seismometers distributed near the epicentre by the Geological Survey of Canada, 24 hours after the 23 November foreshock (magnitude 4.8). About 60 hours later, the main shock of magnitude 6.5 was felt as far south as Washington, DC.

We asked 52 sighting reporters, found through local radio stations and newspapers, to answer a five-page questionnaire. We found three types of luminescence: sparkings without accompanying sound; diffuse light, such as heat-lightning intensity at sunset or sunrise illuminating a good part of the sky; and vertical and

horizontal aurora-like stripes. The first category of EQLs was more closely related to the atmosphere than to the Earth's surface. Fireballs a few metres in diameter often popped out of the ground in a repetitive manner at distances of up to only a few metres away from the observers. Others were seen several hundred metres up in the sky, stationary or moving. Some observers described dripping luminescence droplets, rapidly disappearing a few metres under the stationary fireballs. Only two fire-tongues on the ground were reported, one on snow and the other on a paved parking space without any apparent surface fissure. The colours most often identified were orange, yellow, white and green. Some luminosities lasted up to 12 min. Out of the 46 EQLs reported, 70% were observed 4–35 km north of the epicentral area, reflecting the geographical distribution of the main cities in the region.

Regardless of the fact that before the 23 November earthquake, no seismic activity greater than 1.5 magnitude had been detected by the nearest permanent Charlevoix seismic-monitoring station of the Geological Survey of Canada Network (100 km to the southeast), erratic EQL sightings were reported as early as 1 November. Of the 46 EQLs reported from then until the end of January 1989, 39 (84%) were seen between 1 and 27 November (Fig. 2). During that month, 2 main shocks, followed by 31 aftershocks, with a cumulative magnitude sum of 66.5, were recorded out of a total of 54 ($\Sigma M = 83.7$) until the end of January 1989. The greatest daily number of EQL sightings is highly synchronous with the timing of the two main

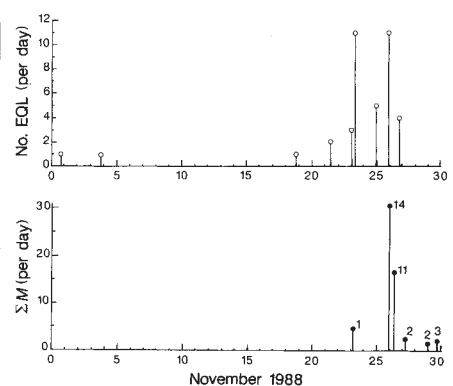


FIG. 2 November mean daily temporal distributions of the number of EQLs reported per day (○) (top) and the sum of the daily magnitudes of seismic events (bottom). ●, No. of seismic events per day.

shocks. On 23 and 25 November, respectively, at 0412 and 1846 (EST), 11 sightings were reported. No seismic activity was recorded on the 24 November, but sightings were observed during the evening of the 21 until the evening of the 26 November. Excluding the foreshock, during which 8 of the 11 EQLs were witnessed simultaneously with the seismic event at 0420, 93% of all luminescences in November were recorded between 1800 and 2400. Because only a few people are awake after midnight, and events are in any case likely to be underreported, the total number of EQL events was probably substantially underestimated. On the other hand, all seismic events of magnitude greater than 0.1 are detected by the seismometer network. These discrepancies make statistical analysis of the two phenomena difficult. Nevertheless, five shocks ($M = 0.8–3.6$; $\Sigma = 10.1$) were registered between 16 and 22 January 1989 and, unlike the three EQLs of December, they were well synchronized with the six EQLs of 16, 19 and 21 January.

Several suggestions have been made^{4–6} to explain these luminous seismotectonic electromagnetic emissions, but none is fully satisfactory. The roles of many physical and chemical factors associated with each seismic event still remain to be elucidated. Furthermore, the inherent difficulties associated with the gathering of controllable field information on such rare phenomena limit the progress of research.

MARCEL OUELLET

Quebec University,
INRS-Eau, PO Box 7500,
Ste-Foy, Quebec,
Canada G1V 4C7

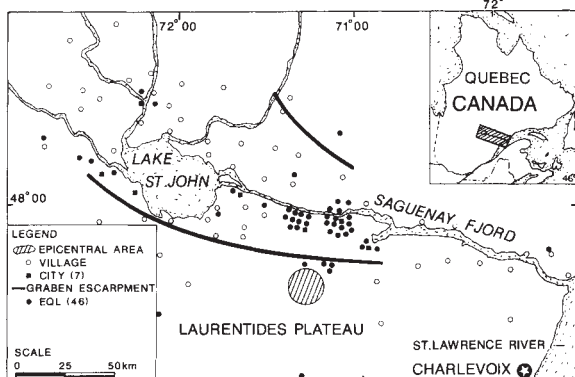


FIG. 1 The Saguenay-Lake Saint John region, showing the epicentral area of the November 1988 earthquakes as well as the EQL sightings and population spatial distributions.

1. Montandon, F. *Geogra. Helvetica (Bern)* **111**(2), 157–178 (1948).
2. Derr, J.S. & Persinger, M.A. *J. sci. Explor.* (in the press).
3. Duberger, R. *et al. Tectonophysics* (in the press).
4. Yamada, I., Masuda, K. & Mizutani, H. *Phys. Earth planet. Int.* **57**, 157–168 (1989).
5. Mizutani, H., Ishido, T., Yokokura, T. & Ohnishi, S. *Geophys. res. Lett.* **3**, 365–368 (1976).
6. Zlotnicki, J. & Le Mouél, J.L. *Nature* **343**, 633–635 (1990).

Dopamine and schizophrenia

SIR—The authors of the papers on the cloning of the D_1 dopamine receptor that recently appeared in *Nature*¹⁻³ suggested that alterations in the number or activity of opposing D_1 and D_2 receptors may be a contributory factor in Parkinson's disease and schizophrenia in that D_1 receptors stimulate and D_2 receptors inhibit adenylyl cyclase. Seeman *et al.*⁴ have demonstrated a link between D_1 and D_2 receptors located on the same postsynaptic membrane, showing that binding of a D_2 agonist to striatal tissues is affected by pretreatment with D_1 -selective antagonist and vice versa, and suggesting cross talk between the receptors mediated by G-protein interactions. This link is missing in more than half the postmortem striata from brains of people with schizophrenia and Huntington disease but is present in control striata¹, indicating a molecular mechanism for the disease state.

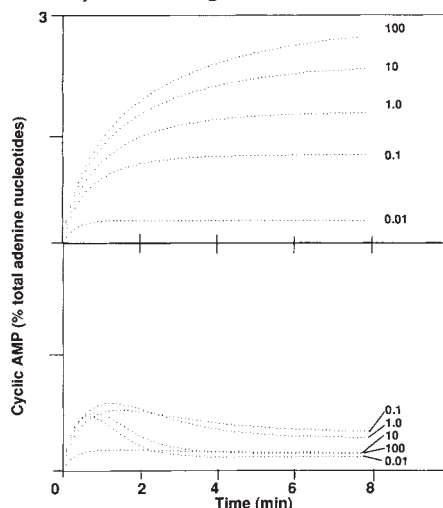
The model is simple and attractive. It can be extended because the existence of separate stimulatory and inhibitory receptors for the same agonist in the same membrane has predictable consequences on cyclic AMP formation. I recently reported^{5,6} a description of prostaglandin-regulated cyclic AMP metabolism in blood platelets based on this model. I suggested that a distinct inhibitory receptor represents a general mechanism of

autacoid (or neurotransmitter) regulation, buffering cellular responses against the transient localized increases in agonist concentration that occur when agonists are produced close to their sites of action⁶. The figure shows a computer simulation of a dose-dependent rise in cyclic AMP with (bottom) and without (top) a slow-acting inhibitory receptor that binds agonist more weakly than the stimulatory receptor.

I propose that D_1 and D_2 receptors together regulate adenylyl cyclase activity on the postsynaptic membrane so that cyclic AMP in neurons of normal individuals is maintained at an essentially constant level despite sharp fluctuations in agonist concentration that must occur when neurotransmitter is released into the tiny volume of the synaptic cleft. Interference with this protective buffering mechanism, which may occur through a reduction or absence of the inhibitory receptor, would result in wild swings in the level of cyclic AMP, reflecting transient but physiological changes in neurotransmitter concentration. The effect of loss of control of dopamine-regulated cyclic AMP level may affect neuronal firing, perhaps explaining manifestations such as those observed in people with schizophrenia.

BARRIE ASHBY

Department of Pharmacology,
Temple University School of Medicine,
Philadelphia,
Pennsylvania 19140, USA



Computer simulations of steady-state cyclic AMP levels in intact cells based on actual data obtained for prostaglandin regulation of cyclic AMP metabolism in blood platelets^{5,6}. Top, The steady-state level of cyclic AMP in the absence of an inhibitory receptor is determined by receptor-linked adenylyl cyclase activity, which increases as a saturable function of agonist concentration (numbers on right) and phosphodiesterase activity which is constant. Bottom, The inhibitory receptor, which converts adenylyl cyclase to an inactive form, binds agonist more weakly than the stimulatory receptor so the concentration dependence of inhibition differs from that for stimulation.

Red herring

SIR—I suspect that the claim by Fry and Bonnet-Bidaud (*Nature* 347, 625; 1990) that an entry in the Han dynasty record book has a bearing on the putative ancient red colour of Sirius is a red herring. The textual statement "At East there is big star called Wolf-horn changes colour" has a sense of immediacy of the sort probably arising from noticing that Sirius was chromatically scintillating with unusual violence as it rose in the east. This relatively unremarkable observation could well produce the commonplace and secure prophecy of "many thieves, robbers". A major and permanent change in colour of the brightest star in the sky would be more likely to imbue the astrologer with the courage to predict failures of harvests or falls of emperors.

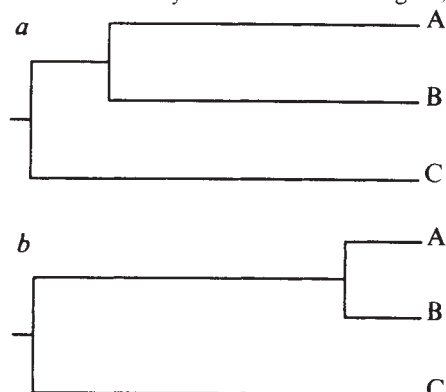
BRIAN WARNER

Department of Astronomy,
University of Texas,
Austin,
Texas 78712, USA

Equal animals

SIR—May in *News and Views*¹ discusses the question of whether all species should be equally valued from the perspective of conservation biology. He and others^{2,3} have concluded that some animals must be considered more equal than others⁴ when assigning spaces in a shrinking — if not sinking — ark.

But how can one assess the relative worth of, for example, two different species of lemur? If diversity is the goal, the key seems to lie in taxonomy, with species for which no close relatives exist receiving greater weight. May described several weighting schemes based on the branching order, or topology, of an evolutionary tree relating the species in question. The length of the branches of such a tree, corresponding to degrees of change between evolutionary diversions, should also be taken into account. The two evolutionary trees shown in the figure,



for example, are topologically identical. Yet the three species in tree *a*, having diverged at nearly the same time, should receive roughly equal weights, while in tree *b*, species C should count almost as much as species A and B combined.

This intuition was first given rigorous treatment by Felsenstein^{5,6} in the context of phylogenetic reconstruction. More recently, we have argued⁷ that Felsenstein's formulation has general applicability to the problem of assessing correlated data related by an evolutionary tree. Conservation biology appears to be an additional field to which these ideas apply.

Felsenstein's weights can be summarized by imagining an evolutionary tree as being constructed of resistant wire. If each leaf is 'grounded', and one volt is applied to the node representing the common ancestor, then the current flowing out of a leaf of the tree represents the relative contribution the corresponding species makes to the total diversity^{5,7}.

These weights are meaningful only in the context of the complete tree, and change if any leaves are pruned. One cannot assign to the various species fixed values which can be added to assess a sub-collection. Instead, if one must choose among several subtrees, the one with the

greatest net conductivity maximizes this measure of diversity.

STEPHEN F. ALTSCHUL
DAVID J. LIPMAN

National Center for Biotechnology
Information,
National Library of Medicine,
National Institutes of Health,
Bethesda, Maryland 20894, USA

1. May, R. M. *Nature* **347**, 129–130 (1990).
2. Atkinson, I. in *Conservation for the Twenty-First Century* (eds Western, D. & Pearl, M.) 54–69 (Oxford University Press, 1989).
3. Vane-Wright, R. I., Humphries, C. J. & Williams, P. H. *Biological Conservation* **55**, 235–254 (1991).
4. Orwell, G. *Animal Farm* (Secker & Warburg, London, 1945).
5. Felsenstein, J. *Am. J. hum. Genet.* **25**, 471–492 (1973).
6. Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).
7. Altschul, S. F., Carroll, R. J. & Lipman, D. J. *J. molec. Biol.* **207**, 647–653 (1989).

Striate cortex periodicity

SIR—Cytochrome oxidase blobs form a regular repeating lattice in the primate striate cortex (area 17) directly related to its functional architecture¹. Recently, Kuljis and Rakic² reported the size of these blobs following bilateral ablation of the retina on the 81st and 120th (E81 and E120) days of gestation. As early bilateral enucleation also reduces the size of striate cortex^{2,3}, this raises the question of whether the periodicity of the cytochrome oxidase lattice is also modified by this manipulation.

Kuljis and Rakic report² that the size of cytochrome oxidase blobs is unaffected by bilateral enucleation, confirming our previous findings¹. We have measured the surface area of area 17 following enucleation at E59, E68, E77, E81 and E110. The extent of the reduction depends critically on the age at which the ablation is performed. At the ages examined by Kuljis and Rakic², enucleation leads to a reduction of striate cortex of only 14–40%, whereas enucleation at E59 and E68 leads

to an areal reduction of more than 70%. We have therefore examined the distribution of blobs following bilateral enucleation at E68.

The mean blob separation in the normal neonate corresponds to published values in the adult⁴ and is similar to that found in the neonate enucleated at E110 (normal, 560 μ m; E110, 568 μ m) and slightly larger than that in the neonate enucleated at E68 (514 μ m; see figure). Although small, the difference between the E68 and the normal interblob spacing was statistically significant ($P < 0.005$). The 8% linear reduction of interblob spacing following enucleation at E68 would correspond to an areal reduction of striate cortex of around 15%, considerably less than that which we observe in this neonate and possibly resulting from the cell atrophy following this operation⁵.

The finding that the reduction in the dimensions of area 17 following early enucleation is not accompanied by an equivalent reduction in mean blob separation is relevant to current theories of cortical parcellation during corticogenesis and the potential role of the sensory periphery. It seems that two levels of specification need to be distinguished. The first occurs early in development, is critically dependent on the presence of the

sensory periphery and determines the areal extent of area 17, at least partly by modulating the levels of cell death and proliferation. The second is apparently independent of the presence of the retinae, operates after the determination of areal borders and specifies the periodicity of cytochrome oxidase blobs.

HENRY KENNEDY
COLETTE DEHAY

Vision et Motricité,
INSERM U94,
16 Avenue Doyen Lépine,
69500 Bron,
France

GWYNN HORSBURGH

Department of Neurobiology,
Anatomy and Cell Science,
University of Pittsburgh Medical School,
3550 Terrace Street,
Pittsburgh,
Pennsylvania 15261, USA

1. Martin, K.A.C. *Trends Neurosci.* **11**, 380–387 (1988).
2. Kuljis, R.O. & Rakic, P. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5303–5306 (1990).
3. Dehay, C., Horsburgh, G., Berland, M., Killackey, H. & Kennedy, H. *Nature* **337**, 265–267 (1989).
4. Tusk, T.C., Kaboreed, W.S. & Wong-Riley, M.T.T. *Visual Neurosci.* **4**, 185–204 (1990).
5. Dehay, C., Horsburgh, G., Berland, M., Killackey, H. & Kennedy, H. *J. Eur. Neurosci.*

Forlorn hope for malaria vaccine?

SIR—Successful vaccination against viruses and bacteria has, it seems, led to the notion that vaccines could also be developed against malaria parasites. But malaria parasites reproduce sexually within the insect host. Sexual reproduction leads to the maintenance of and increase in heterozygosity, producing phenotypes ready to cope with almost any change that we may impose^{1,2}. This important difference seems either to have been misunderstood or disregarded, and I wonder if the notion of finding a valid vaccine against malaria parasites is not flawed from the outset.

It is, of course, true that people living in regions where malaria is endemic develop an immunity, at least to the severe and dangerous symptoms of the disease. But, as anyone who has worked in such places well knows, if these immune people move away from the region, some of them contract malaria after two weeks or so. They have not lost the immunity to their 'home' parasites — although they will do so unless constantly challenged by them — they are being affected by different local phenotypes of *Plasmodium*.

It is this neglect or misunderstanding of the consequences of sexual reproduction, and the maintenance and constant enrichment of genetic diversity that is its corollary, that has in my opinion led the well-intentioned malaria vaccine lobby astray. The same may well apply to the search for

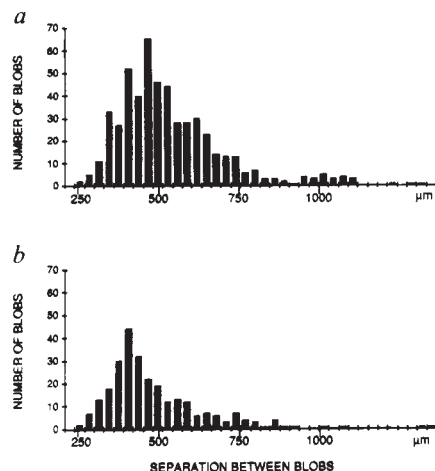
a vaccine to combat other eukaryote parasites such as *Trypanosoma*, complicated even further with its sequential antigenic variation, and *Schistosoma*. One witnesses a similar flaw in recent approaches to find a vaccine to combat east coast fever (ECF) in cattle. The argument goes like this: an effective vaccinia virus recombinant vaccine has been developed against rinderpest and "this encourages us to pursue similar strategies in developing vaccinia for other livestock diseases [that is, ECF]", to cite ref. 3. Rinderpest and ECF are, to be sure, both livestock diseases but the former, like yellow fever, is a virus disease and the latter, like malaria, is a protozoal disease.

Perhaps the enormous and expensive effort now going into a search for a malaria vaccine would be better spent on a search for new methods to prevent contact between insect and man⁴.

J. D. GILLET

Department of Medical Parasitology,
London School of Hygiene and
Tropical Medicine,
Keppel Street,
London WC1E 7HT, UK

1. Bell, G. *The Masterpiece of Nature* (Croom Helm, London, 1982).
2. Lively, C.M., Craddock, C. & Vrijenhoek, R.C. *Nature* **344**, 864–866 (1990).
3. Ole-Moi Yoi, O.K., Nayar, A., Iams, K., Musoke, A.J. & Yilma, T. *Ann. N.Y. Acad. Sci.* **569**, 174–182 (1989).
4. Gillett, J.D. *Trans. R. Soc. trop. Med. Hyg.* **79**, 12–20 (1985).



Frequency distribution histogram of adjacent blob separation on parasagittal sections for a normal neonate (a) and a neonate which underwent bilateral enucleation at E68 (b).

Signature of deadly water

Eric Ashby

A Science of Impurity. By Christopher Hamlin. *Hilger*: 1990. Pp. 342. £45, \$39.95.
Water, Engineering and Landscape. Edited by Denis Cosgrove and Geoffrey Petts.
Belhaven: 1990. Pp. 214. £30, \$49.

IN the year 1778 a device was patented by Joseph Bramah which brought death to multitudes of people. It was not a weapon of war, it was the water closet. The immediate causes of death were epidemics of cholera and typhoid, and it was not until a century later that the link between Bramah's device and these epidemics was fully confirmed. The original invention of the water closet is attributed to Sir John Harington, in 1596, but it was Bramah who manufactured and installed them in private homes.

The consequences for city dwellers are familiar. Human excreta could no longer be carried away as night-soil from cess-pools. The privy discharged into a sewer; the sewers discharged into rivers; private companies drew water from the rivers and returned it to the taps of private houses. As early as 1828 William Lambe, a distinguished physician, was warning the public that drinking water, known to contain "the decayed and decaying remains of myriads of animals and vegetables, in every stage of decomposition and putrefaction" might be harmful to health. Analysts failed to find anything in tap water to justify this warning; to which Lambe replied: "It is . . . vain to say, that where nothing is discovered there is nothing wrong."

This is the dilemma that haunted sanitarians throughout the nineteenth century. So much has been written about the history of public health in Britain and America that one is tempted to wonder whether there is anything new to be said about it. Hamlin himself in *A Science of Impurity* adds a "bibliographic essay" of 23 pages: an annotated list of books and papers recording an enormous amount of information and opinion. It would seem audacious, not to say rash, to write anything more.

Nevertheless, Hamlin has used this mass of material to create a fresh theme. It is a history of the 'experts' who analysed water in Britain from 1800 to 1898, and the part they played in shaping public policy. For the reader who thinks he is already familiar with the story the book has some nice surprises. Is it, for instance, well known that the first water analysts were not concerned with the pollution of water or its danger to health? They were concerned with its enrichment by salts which were believed to cure dyspepsia, rheumatism and other disorders; for there was money to be made out of this. Bath and

Harrogate, by virtue of their mineral waters, became spas frequented by wealthy patrons. Their success prompted other towns to promote the therapeutic value of their springs. To this end, consultants were employed to analyse their water (mainly for carbonates, sulphates and chlorides) knowing that a favourable analysis (whatever that might mean) would

be welcomed by the civic authorities and presumably by local medical practitioners.

It was a natural extension of this interest in analysis for chemists to be willing to act as consultants for the private water companies that supplied London and other cities with water. At the same time, sanitarians began to realize that there might be ominous organic impurities in water that would not be disclosed by the chemists. Discussions took on a flavour of adversarial piquancy. On one hand chemists, looking only for inorganic salts, found nothing in London's water inimical to health; on the other hand there were sanitary engineers who took the view (to quote one of them):

... a stream which receives daily the evacuations of a million human beings . . . with all the filth and refuse of various offensive manufacturers . . . cannot require to be analysed, except by a lunatic, to determine whether it ought to be pumped up as a beverage for the inhabitants of the Metropolis of the British empire.

The controversy became more heated in 1850 when Arthur Hassall published *A*

Microscopical Examination of the Water Supplied to the Inhabitants of London and Suburban Districts, with pictures of the animalculae teeming in samples of tap water, only months after cholera had killed 14,000 people in Greater London; to which the retort of chemists was that to swallow animalculae was "no more harmful than eating fish". But still nothing was isolated from the water that could account for cholera or typhoid.

It was not until the 1880s that germs replaced miasmas as agents for disease. But there was no dramatic breakthrough in the treatment and analysis of water. Hamlin's important contribution in this book is to demonstrate that there was no simple "sequence of discovery, recognition of implications, and action." Of

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Water closets — brought death to multitudes of people.

course bacteria could not be counted or identified from water samples, and even when plated out on gelatin, pathogens could not at first be distinguished from ubiquitous colonies of coliforms. "Analysts", concludes Hamlin, "never did find the simple and readily discoverable signature of deadly water." So for years the adversarial engagement stumbled on. Reformers invoked science to sanction change; opponents of reform invoked science to oppose change. Science provided convenient answers for both sides.

How familiar this is in 1990, no longer for water but for other environmental dilemmas. And, in a way, even for water; for there are still overtones of the Victorian concern for water to be "clear, wholesome, and palatable", as I realize when I purchase a two-litre bottle of mineral water at the supermarket on the grounds that my tap water is deficient in the third of these qualities.

Hamlin has written an important book, excellently documented in clear English. The narrative is in places very detailed, but the reader is led carefully through the

Mary Evans

maze of claim and counter-claim, and emerges at the end with some fresh insight into a fascinating episode in British social history.

From microparasites that pollute water to the macroparasite, Man, who exploits waters: in 1988 geographers at Loughborough University met in a series of seminars to discuss the impact of water engineers on the environment. *Water, Engineering and Landscape* contains the papers given at these seminars. The topics include hydrology in sixteenth-century Venice, the draining of the East Anglian fenlands in the seventeenth century, and the extraordinary and little-known campaign conducted by François Roudaire in the nineteenth century, to create "*la mer intérieure*" in vast stretches of the Sahara which lie below sea level, by building canals from the gulf of Gabes to admit the waters of the Mediterranean. It was a grandiloquent gesture of French imperialism which created a storm of controversy in the 1870s. In addition to these essays in historical geography there are

papers on contemporary problems: hydro-technology in Quebec, the fiercely disputed project for a dam on the Danube, and the perennial problem, inseparable from the aspirations of Welsh nationalists, of how Wales should deploy its rich resources of water.

It was doubtless the repertoire of the Loughborough geographers that determined this choice of topics. The value of a seminar lies in the interaction between those who take part in it rather than in the content of individual papers. The papers are interesting but one does not get the impression that any one of them is relevant to any other. To use the analogy of music, Hamlin's book has the coherence of a symphony (even including some of the *longueurs* that occasionally afflicted Schubert); whereas the book edited by Cosgrove and Petts is more like a Victorian *salon* concert, where each of the 11 artists performs his party piece. □

Eric Ashby is at Clare College, Cambridge, CB2 1TL, UK.

Technical targets

John Heilbron

Science, Technology, and Reparations: Exploitation and Plunder in Postwar Germany. By John Gimbel. Stanford University Press: 1990. Pp. 280. \$29.50.

As usual, a Hague Convention applied. It allowed seizure of enemy property that had direct military use or, in the case of occupation, that belonged to the state. And, as usual, the victors ignored the convention. That was a boon for many American industrialists and also for John Gimbel, whose latest book on the occupation of Germany concerns the 'exploitation' (the term of art used by the Americans) of German industry for private benefit in the United States.

The exploitation developed from wartime intelligence missions, such as ALSOS (which rounded up German nuclear physicists and their 'secrets'), BIOS and CIOS (respectively, the British and the Combined Intelligence Objectives Subcommittee). The main charge of these missions was to acquire personnel and *matériel* that might help in the war against Japan and, more generally, to insure that the Allies would have the advantage of all significant military developments made in Germany during the war. In furtherance of this programme, the Americans corralled specialists in military hardware in detention centres coyly called 'Ashcan' and 'Dustbin' and brought rocket scientists and their hardware to the United States through project 'Overcast'.

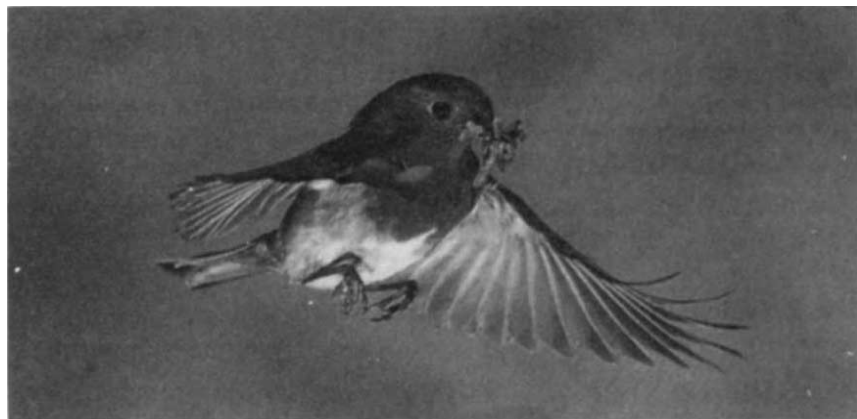
Overcast was the beginning of a more general importation of German brains, the notorious 'Paperclip', which aimed not only at exploitation but also at 'denial'. A persuasive argument for legitimizing Paperclip in the United States was the importance of denying the other occupying powers the opportunity to recruit or kidnap German technical experts living in the American Zone.

A draft press release, dated 11 March 1946, laid out the rationale for Paperclip so clearly that the Joint Chiefs of Staff classified it 'secret'. Gimbel prints it in full. It states Paperclip's goal as 'the complete exploitation of Germany for technical information'; provides for a limited importation of German specialists, all clean politically and intellectually "comparable to Prof. Einstein"; stipulates exploitation for civilian-industrial as well as military-technological purposes; and characterizes its way of business as

"vacuum cleaner methods".

The most energetic agency for industrial exploitation was the Office of Technical Services (OTS) of the Department of Commerce. Its director, John C. Green, liked to say that German trade secrets were "the only solid and permanent reparations we are going to get out of this war". The OTS recruited men from American industries to inspect plants dedicated to similar businesses in Germany, to identify useful machinery for removal to the United States, to interrogate technicians and managers, and to examine blueprints, drawings and manufacturing records. Their activities abroad were coordinated by the Field Intelligence Agency, Technical (FIAT) which also arranged for the wholesale copying of the records of German firms. The reports of the inspection teams and FIAT's micro-filmed records were sold for the cost of reproduction by OTS. (Amtorg, the Soviet Purchasing Agency in the United States, was OTS's most faithful customer.) No industry, however humble, escaped the vacuum cleaners of FIAT and OTS. There were obvious targets like synthetic fuels, artificial rubber, tape recorders, machine tools, optical apparatus and jet planes, but also manure spreaders, cosmetics, milk cans and stuffed animals. The means of making teddy-bears, a German speciality, stood revealed to the world.

There was not much chance that German industry would recover in an environment that threatened teddy-bears. The United States' Office of Military Government (OMGUS), commanded by General Lucius D. Clay, decided that it could not fulfil its mission with FIAT in the field. Gimbel tells the story of the resultant interagency battle clearly and economically. Green won the opening round by coaxing President Truman to write to OMGUS about the importance of FIAT. Clay retaliated by asking that an exact account be made of the dollar value of the technical know-how and equipment



Avian aerodynamics — A robin approaching the nest with its prey. *Bird Flight* by Robert Burton discusses some of the skills involved in getting, and staying, airborne. Published by Facts on File, price is £14.95. □

removed as 'reparations'. No one knew how to do that. Also the British, who had finished most of their exploiting, insisted that the Americans cease too as a condition for joint governance of their newly created Bizone. Eventually, over Green's protests ("it will be a national tragedy . . . if we allow the doors to shut before we have added all of the best of Germany's technical knowledge to our own"), FIAT was terminated on 30 June 1947.

Gimbel's story invites two sorts of reckonings. What was the value of the intellectual reparations? Where does the ethical balance lie? Molotov made the first reckoning. At a meeting of the Council of Foreign Ministers early in 1947, he accused the United States of removing assets and information worth ten billion dollars. The Secretary of State, George Marshall, indignantly countered with a sum a thousand times smaller. A few days after the exchange, the United States' government tried to collect estimates from the War, Navy and Commerce departments. No luck. OMGUS urged the various German states under its control to obtain estimates from the affected firms. No luck. The firms feared that further espionage or greater tax liability would follow disclosure. Gimbel has searched archives at home and abroad for the answer. From a few hard numbers and some anecdotal information, he concludes that Molotov was not far wrong.

Gimbel offers no explicit judgment about the right and wrong way of the exploitation. His subtitle, however, as well as his recital of the special favours and profits accorded domestic firms by OTS and the military, suggest where he stands. Certainly, neither the favouritism at home nor the race among the occupying powers abroad to seize German technical information, plant and personnel makes an edifying story. But it is hard to condemn the exploitation in principle. Why should the defeated enemy alone enjoy the fruits of the industrial research stimulated or commissioned by the Nazi war machine? FIAT did not come to an end because Clay occupied a higher ethical ground but because exploitation was proving counter-productive to OMGUS' mission to restore German industry to a level sufficient to supply the needs of the Bizone. In the final reckoning, the \$10 billion was insignificant in comparison with the cost of maintaining a dependent Germany.

Gimbel's detailed report from the archives brings valuable new information; an example of a productive interagency dispute; and, above all, an opportunity to think through again the relations among the victors and the vanquished just after the second, and, we trust, the last world war. □

John Heilbron is at the Office for History of Science and Technology, University of California, Berkeley, California 94720, USA.

Community spirit

Geoffrey Fryer

Freshwater Ecology. By Michael Jefferies and Derek Mills. *Belhaven*: 1990. Pp.285. Hbk £27.50, \$49; pbk £12.95, \$15.

As the writers of this book on freshwater ecology tell us, its main theme "is to explore the processes that sculpt and develop whole communities and what happens when human activity intervenes." By way of introduction there are elementary accounts of the properties of water, the hydrological cycle and related matters, and chapters on standing and flowing waters and plant life. All this is

with a glance at the future. Coverage of these applied aspects distinguishes this book from most other texts on freshwater ecology. Treatment is, however, rather elementary and the promise of an analysis of what happens to communities when man interferes is fulfilled to only a limited extent. The general effects of the various stresses considered are indeed often depressingly predictable. On the other hand, to work out the details of how a community responds to environmental changes is a complicated business that demands more knowledge of how it was initially organized than is often available, as well as precise understanding of the factors that stress it. Detailed treatment is not accorded here. This is not to say that these sections will not be helpful to undergraduates, senior school pupils and those



Visible pollution — man's interference with natural waters.

straightforward and covers well-trodden ground. It leads up to a consideration of animal communities that are looked at from the standpoints of abiotic influences and biotic interactions; matters that constitute one of the avowed aims of the work. The former includes a welcome, but all too brief, reference to the often neglected influences of historical and geographical factors on the make-up of communities. Making use of up-to-date examples of manipulation experiments as well as of observational investigations, the section on biotic interactions is certainly the most interesting thus far, but it and the preceding section are allocated only 47 pages, including illustrations, in which to cover this enormous and complex topic.

The rest, slightly less than half the text, is concerned with man's interference with natural waters and deals with eutrophication, the current 'hot' topic of acidification, the effects of land use, fish farming, water abstraction and transfer and hydroelectric schemes, and concludes

concerned with conservation and water management who require a broad overview of these matters, who appear to be the most likely readers.

The style is generally clear, but to use 'plankton' and 'benthos' as plural nouns is objectionable. Errors are not rare. For example, 'blanket weed' is *Cladophora*, not *Enteromorpha*, which is very different; not all aquatic plants "take nutrients largely from the substrate", *Daphnia obtusa* is quite the opposite of a large-lake species, the present Lake Victoria does not date from Miocene times but is much younger, nor could the Nile perch have arrived there via the turbines of the Owen Falls Dam as it did not occur in the river below the impoundment. Helpful and attractive, but poorly reproduced, photographs enliven the text. The line drawings tend to be crude. □

Geoffrey Fryer is at the Institute of Environmental and Biological Sciences, University of Lancaster, Lancaster LA1 4YQ, UK.

Science Photo Library

Success story

J. D. Ross

Seed Dormancy in Grasses. By G. M. Simpson. Cambridge University Press: 1990. Pp. 297. £35, \$54.50.

If we are what we eat then, indeed, all flesh is grass. From the earliest days of agricultural society man has depended on the grass family for his bread, beer and beef. The Gramineae is one of the largest plant families with more than 600 genera and perhaps 10,000 species, found in virtually all ecosystems (including marine and Antarctic) and covering a greater percentage of the land surface than any other family. Such is the distinctiveness of grasses that even the botanically uninitiated can easily recognize this group by the characteristic vegetative structure and reduced floral parts. Eight genera contribute the cereal grasses while, on the other hand, seven of the worst weeds in the world are also grasses. There exists a massive scientific literature on the biology of the crops, pasture grasses and weeds — about 350 species in all; of the others, the vast majority, we know very little beyond a rudimentary taxonomic description.

One factor contributing to the natural success of the Graminae is their superb adaptation to being grazed, cut, trampled and burned. Another important feature is their relatively large and aggressive reproductive output; they tend to be prolific in the production of mobile and invasive seeds.

In the cereal growing regions of the world one of the most pernicious weeds is the wild oat, *Avena fatua*; the key to its success arguably being its dormancy mechanism which results in infestations lasting many years, even under the

cleanest of cultivation regimes. G. M. Simpson, the author of this somewhat anachronistic monograph, has spent much of his research career studying this phenomenon. This approach is unusual but successful, for here is a physiologist who knows the taxonomy of the family and sees the usefulness of such knowledge. The deeply researched wild oat is used as a model alongside reviews of the comparable knowledge on other species. Over a thousand references are cited in reviewing endogenous influences on dormancy and interactions with environmental factors. Many of these reports originated from the Weed Research Organization, now regrettably no longer operating, weeds of our crops being perhaps too near the market-place for British science.

As with other physiological systems the

reductionist approach has yielded an impressive amount of information, but in bringing these data together we have failed to elucidate a general understanding. When this happens the timid researcher switches to another problem and the brave one moves into modelling. In the last section of this book there is an attempt to do just that: using Odum-style energy flow diagrams the author draws attention to the holistic nature of the physiology of the dormant seed. All in all, *Seed Dormancy in Grasses* provides a most useful case study with a balanced coverage and an optimistic attitude to the future, the latter being something I did not share until I read this book. □

J. D. Ross is in the Department of Botany, University of Reading, Reading RG6 2AS, UK.

Highly prized molecule

Jim S. Owen

Cholesterol Metabolism, LDL, and the LDL Receptor. By N. B. Myant. Academic: 1990. Pp. 465. £46.95, \$69.95.

THOSE who aspire to a Nobel prize, but who have yet to select their *magnum opus*, would do well to muse over this volume. No fewer than 14 laureates have chosen to study cholesterol, a molecule not only essential as a structural component of eukaryotic membranes and a precursor of steroid hormones and bile acids, but also an aberrant one because of its role in the cause or exacerbation of several diseases, most notably atherosclerosis. Cholesterol homeostasis at the cellular and molecular level is the theme of Myant's book, which in a sense charts the story and sequel of the most recent award, in 1985, to M. S. Brown and J. S. Goldstein who fittingly provide a foreword.

The journey begins with a guided tour through the maze of information on the regulation of HMG-CoA reductase, the intracellular rate-limiting enzyme of cholesterol biosynthesis, at the level of both gene and protein. There is a diversion in the form of the biochemistry of acyl-CoA:cholesterol acyltransferase or ACAT — but only a brief one because this important sensor of a cell's unesterified cholesterol content remains poorly characterized, almost embarrassingly so when considered alongside, as here, the plethora of details on the two other key monitors — before we embark, at an ever increasing pace, into systematic and detailed descriptions of the structure and molecular genetics of the two principal proteins involved in delivering cholesterol to the cell: the low-density lipoprotein (LDL) receptor and its ligand, apoB-100,

the major cholesterol carrier in human plasma. *En route* there are minor detours — to absorb relevant information on such supporting acts as the scavenger receptor, apoB-48, the LDL-receptor-related-protein (LRP), Lp(a) and modified LDL — many of which would constitute major holdups were it possible to include this year's literature. But the final route is not quite circular; homeostasis also demands the ability to return cellular cholesterol to the circulation and neither this, nor the importance of the HDL carrier system to the process, particularly the role of the pre-beta-migrating HDL particles, are discussed.

The last decade has seen remarkable advances in the understanding of cholesterol transport and metabolism. Many illuminate other areas of cellular and molecular biology, including the functioning and recycling of cell-surface receptors, the concept of several different functional domains within a single polypeptide, and the control of expression of genes encoding inducible proteins. This book is an indispensable guide to these achievements and many people, myself included, will be grateful for its timely appearance: the scientific and clinical teachers of lipoprotein metabolism and familial hypercholesterolaemia for its comprehensive reviews of several complex issues; the multidisciplinary army of workers in the field, particularly those without training or experience in molecular biology; and the growing number of interested non-specialists, including, for example, parasitologists who increasingly recognize that the cholesterol of pathogenic organisms is often an essential nutrient derived by receptor-mediated endocytosis of host lipoproteins. The chronicler of the next ten years of cholesterol metabolism would be well advised to begin now. □

Jim S. Owen is in the Academic Department of Medicine, Royal Free Hospital School of Medicine, London NW3 2PF, UK.

New in paperback

■ *International Environmental Diplomacy* edited by John E. Carroll concerns itself with the international policy aspects of environmental problems, their management and resolution. Published by Cambridge University Press, price £12.95, \$24.95

■ From Earthscan comes *Elephants, Economics and Ivory* by Edward Barbier, Joanne Burgess, Timothy Swanson and David Pearce. The authors set out to show how the careful management of elephants as a resource can best serve African interests. Price £8.95.

■ Reprinted by Gordon and Breach is *Lavoisier — The Crucial Year* by Henry Guerlac which examines a key period in the career of Antoine-Laurent Lavoisier, the chief architect of the chemical revolution. Price \$27.00.

■ *Neural Computing: An Introduction* by R. Beale and T. Jackson is aimed at computer science undergraduates and anyone interested in neural networks and artificial intelligence. Starting from basics, it goes on to cover all the most important approaches to the subject. Published by Hilger, price £15.00 □

Orbits of shepherd satellites deduced from the structure of the rings of Uranus

Carl D. Murray & Robert P. Thompson

Astronomy Unit, Queen Mary and Westfield College, Mile End Road, University of London, London E1 4NS, UK

The sharp edges of the rings of Uranus require the presence of small satellites to halt radial spreading. There are several key locations in the ring system where satellites would be most effective at confining rings. Some of these are close to resonances with the moon Cordelia, and there is indirect evidence for the existence of at least two of these satellites.

SINCE their discovery in 1977 the rings of Uranus have posed a number of dynamical problems. The nine rings (6, 5, 4, α , β , η , γ , δ and ϵ) detected by ground-based occultation studies are very narrow (typically <10 km) and some, particularly the outermost ϵ ring, have variable width and precess uniformly^{1,2}. In the absence of an atmosphere, unconfined narrow rings should spread due to the combined effects of particle collisions, Poynting–Robertson light drag and plasma drag, implying ring lifetimes of $\sim 10^7$ yr (ref. 3). Small satellites with resonances at each edge of a narrow ring can, however, provide torques to counteract the spreading and exert a ‘shepherding’ effect on the ring⁴. Voyager 2 images of the Uranus system led to the discovery of ten new satellites⁵ and two of these, Cordelia and Ophelia, were found to have resonances coinciding with inner and outer edges, respectively, of the ϵ ring^{6,7}. Thomas *et al.*⁸ have estimated the dimensions of the new satellites, and their orbital elements have been calculated by Owen and Synnott⁹. Extensive searches for other satellites place an upper limit of radius ~ 10 km on

the size of undetected moons in the vicinity of the ring system⁵. The location of small satellites can, however, be predicted by other means^{10–12}.

The Voyager 2 images in backscattered light showed a faint ring (1986U1R, now renamed the λ ring) at 50,024 km, and another faint ring was detected ~ 75 km outside the β ring. A unique, high-phase-angle image (Fig. 1) using an exposure of 96 s showed the presence of numerous dust rings in addition to the eleven rings visible in backscattered light. Occultation data derived from the Voyager 2 photopolarimeter experiment suggested two more rings and three ring arcs¹³. A more detailed study of the same data revealed 19 anomalous occultation features¹⁴. The discovery of an extended hydrogen exosphere at Uranus¹⁵ has put further constraints on the lifetimes of rings. The drag torque leads to estimates of ring ages ranging from 10^2 – 10^3 yr for the dust rings¹⁵ to 6×10^8 yr for the ϵ ring¹⁶, despite the effects of any shepherding satellites. This may argue in favour of a continual creation of ring material^{16,17}, unless we are living in a special epoch when the rings happen to exist.

Although the problem of ring lifetimes under atmospheric drag remains, of more immediate concern is an explanation for the ring structure visible in the images at low and high phase angles; even if the shepherding process is ineffective over time-scales comparable with the age of the Solar System, the observations of narrow rings with sharp edges have to be accounted for. Only two shepherding satellites have been detected so far and although the locations of other satellites are unknown, a number of techniques can be used to identify likely orbits. We have used the orbits of the known rings and satellites combined with data derived from the image of the dust rings to derive

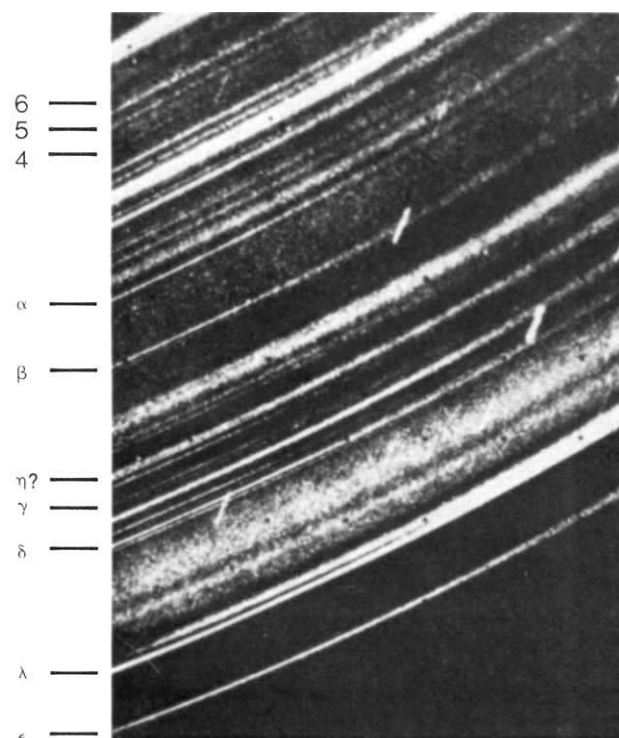


FIG. 1 A Voyager 2 image (FDS 26852.19) of the rings of Uranus taken using a 96-s exposure at a phase angle of 172.49° . The image was targeted to minimize radial smear along the left-hand edge, and covers a radial distance of $\sim 11,000$ km and an azimuthal range of $\sim 10^\circ$. The azimuthal smear due to keplerian motion while the shutter was open varies from $\sim 1.7^\circ$ at the top of the image, to $\sim 1.1^\circ$ at the bottom. The radial locations of 10 of the rings visible in backscattered light are indicated. (Photograph courtesy of NASA/JPL).

TABLE 1 Locations of hypothetical satellites

Satellite	a_s (km)	Resonances					
HS1	41,621	6 129:130	5 45:46	4 29:30			
HS2	42,024	6 150:149	5 133:134	4 51:52			
HS3	42,034	6 143:142	5 140:141	4 52:53			
HS4	42,044	6 136:135	5 147:148	4 53:54			
HS5	42,202	6 77:76	4 76:77	α 11:12			
HS6	42,901	6 27:26	5 43:42	4 87:86			
HS7	44,328	6 12:11	4 76:77	β 22:23			
HS8	44,949	α 132:131	β 42:43	γ 11:12			
HS9	45,059	α 89:88	β 50:51	η 14:15			
HS10	45,158	α 69:68	β 60:61	γ 12:13			
HS11	46,006	α 24:23	β 90:89	η 26:27			
HS12*	46,950	η 139:140	γ 46:47	δ 23:24			
		λ 10:11					
HS13	47,666	β 16:15	η 65:64	δ 50:51			
HS14	47,779	η 53:52	δ 61:62	λ 14:15			
HS15	47,857	η 47:46	γ 139:138	δ 72:73			
HS16*	47,921	η 43:42	γ 109:108	δ 84:85			
		λ 15:16					
HS17*	47,967	α 10:9	γ 94:93	δ 96:97			
HS18*	47,972	β 14:13	γ 93:92	δ 97:98			
HS19*	47,979	γ 91:90	δ 100:101	ϵ 10:11			
HS20	48,044	η 37:36	γ 77:76	δ 126:127			
HS21*	49,630	δ 25:24	λ 84:85	ϵ 22:23			
HS22*	49,750	δ 23:22	λ 121:122	ϵ 24:25			
HS23	50,287	δ 17:16	λ 128:127	ϵ 40:41			
HS24	50,745	γ 11:10	λ 47:46	ϵ 90:91			
HS25	51,461	δ 11:10	λ 24:23	ϵ 121:120			
Ophelia	53,764	γ 6:5	ϵ 14:13				

Locations of hypothetical satellites (HS) that have three or more first-order eccentric resonances within 1 km of ring edges. Inner eccentric resonances are of the form $p+1:p$ whereas outer eccentric resonances are of the form $p:p+1$, where p is an integer. The two known resonances between the satellite Ophelia and ring edges are included for completeness.

* Satellites that lie in gaps in the image of the dust rings. Satellite 22 is within 2 km of the orbit of the known satellite Cordelia.

possible locations of the satellites that constrain the rings.

Cleared regions

A satellite can be indirectly detected in a ring system by the gap that it creates in surrounding ring material due to the presence of a large chaotic zone where particle orbits are scattered. A first-order resonance can occur between a ring particle and a perturbing satellite if the mean motions, or mean angular velocities, n and n' of the particle and the satellite satisfy the relation $(p+1)n' \approx pn$, where p is an integer. Wisdom¹⁸ showed that the overlap of adjacent first-order resonances occurs when $p \geq 0.51 \mu^{-2/7}$ where μ is the ratio of the mass of the perturber to the mass of the central body. For a satellite of density $\rho = 1.5 \text{ g cm}^{-3}$ and radius r (km) orbiting Uranus, $\mu \approx 7.2 \times 10^{-14} r^3$. For Cordelia $r = 13 \pm 2 \text{ km}$ (ref. 8) and the overlap zone will occur for $p \geq 321$ corresponding to a predicted width of $206 \pm 27 \text{ km}$ for a gap centred on the satellite's semimajor axis of $a_c = 49,752 \text{ km}$. Using the same assumed density, the radius of a satellite capable of generating a gap of width W at semimajor axis a is given by

$$r = 7.8 \times 10^3 (W/a)^{7/6} \text{ km} \quad (1)$$

Although there is no optically thick ring material within $\sim 1,200 \text{ km}$ of Cordelia, the dust rings seen in the forward-scattered image (Fig. 1) provide a useful means of detecting the presence of small satellites. Using the left-hand section of the image to reduce the effects of radial smear and binning pixels to improve the signal-to-noise ratio, we have calculated the radial location of all ring features and gaps in the image¹⁹⁻²¹ with a typical accuracy of $\pm 7 \text{ km}$ (Fig. 2a, b); the narrow range of azimuth precludes the possibility of calculating semimajor axes and eccentricities from the image. Similarly, arcs of material with azimuthal ranges greater than that covered in the image

TABLE 2 Possible resonances

Satellite	a_s (km)	Resonance	Type	a_{res} (km)	$a_{res} - a_s$ (km)
HS6	42,901	Cordelia 5:4	corotation	42,889.6	-11.4
		Cordelia 5:4	eccentric	42,895.0	-6.0
HS11	46,006	Cordelia 9:8	corotation	46,002.9	-3.1
		Cordelia 9:8	eccentric	46,004.4	-1.6
HS12*	46,950	Cordelia 12:11	corotation	46,954.1	+4.1
		Cordelia 12:11	eccentric	46,954.9	+4.9
HS13	47,666	Cordelia 16:15	corotation	47,661.2	-4.8
		Cordelia 16:15	eccentric	47,661.7	-4.3
HS14	47,779	Cordelia 17:16	corotation	47,785.4	+6.4
		Cordelia 17:16	eccentric	47,785.8	+6.8
HS15	47,857	Desdemona 3:2	eccentric	47,851.5	-5.5

Possible resonances between the hypothetical satellites listed in Table 1 and known satellites in the Uranus system. All first-order resonances of the fifteen known satellites that lie within 12 km of a hypothetical satellite in the range $38,000 \leq a_s \leq 48,500 \text{ km}$ are included.

* Satellites that lie in gaps in the image of the dust ring.

cannot be detected. The radius of the centre of each of the rings (6, 5, 4, α , β , η , γ , δ , λ and ϵ) at the time the image was taken is indicated. All these rings have counterparts in the forward-scattered image, with the possible exception of the η ring. The δ ring was incorrectly identified in Fig. 13 of ref. 5.

Figure 2b shows a gap of width $261 \pm 14 \text{ km}$ centred at an orbital radius of $49,724 \text{ km}$ in good agreement with Cordelia's orbit, given the uncertainty of $2a_c e_c - 47 \text{ km}$ (ref. 9) introduced by Cordelia's eccentricity. There are at least five other regions that appear to be completely empty: (1) a 137-km gap at $46,962 \text{ km}$; (2) a 100-km gap at $47,315 \text{ km}$; (3) a 185-km gap at $47,508 \text{ km}$; (4) a 151-km gap at $47,955 \text{ km}$; (5) a 75-km gap at $49,955 \text{ km}$. Although the positional errors of the gaps are $\pm 7 \text{ km}$, uncertainties in satellite eccentricities introduce a further error of $\sim 20 \text{ km}$. Assuming each gap is the result of perturbations by a single satellite, equation (1) implies satellite radii of 9 km, 6 km, 12 km, 9 km and 4 km, respectively, at the above locations. The outermost gap just inside the λ ring has a badly removed rescan mark on the image giving the illusion of material in the scan data (Fig. 2b). There are other low-intensity regions in the image but we cannot make definitive statements about other gaps until more attention has been given to the removal of the 'dark current' in the image, a necessary process for any detailed photometric analysis²². Thus, the 'ramp' visible in Fig. 2a may disguise the existence of other gaps. Although our data seem to confirm the existence of a ring just beyond the β ring, we found no evidence for the other ring features detected in the Voyager occultation data^{13,14} which lie in the radial range covered by this image.

We have made an exhaustive search for first- and second-order $((p+2)n' \approx pn)$ resonances between the edges of rings visible in the forward-scattered image and the fifteen known satellites. No such resonances were found apart from those already discovered by Porco and Goldreich⁵ involving the satellites Cordelia and Ophelia and the ϵ , δ , λ and γ rings, and those that might be expected by chance due to the large number of first-order Cordelia resonances in the region $48,500 \leq a \leq 51,000 \text{ km}$. The known satellites alone cannot account for the structure of the rings of Uranus.

Satellites and resonances

A narrow ring can be prevented from spreading by the combined perturbing effects of two satellites, one inside the ring with a resonance at the ring's inner edge and one outside the ring with a resonance at the outer edge. If we consider first-order resonances, $(p+1)n' \approx pn$, then two terms dominate the Fourier series expansion of the perturbing potential. The angles associated with these terms, known as the resonant arguments, are

$$\begin{aligned} \theta_1 &= (p+1)\lambda' - p\lambda - \tilde{\omega} \\ \theta_2 &= (p+1)\lambda' - p\lambda - \tilde{\omega}' \end{aligned} \quad (2)$$

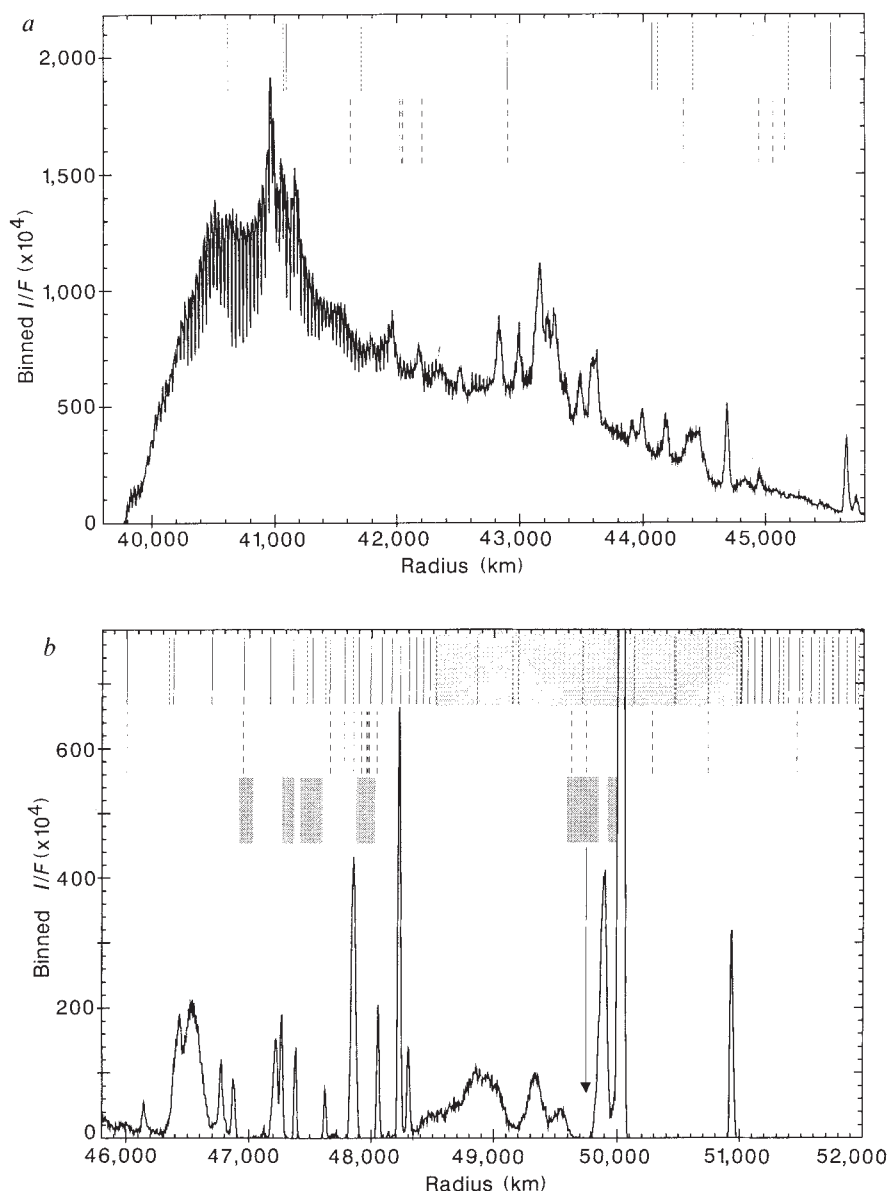
where λ and ω denote the mean longitude and longitude of pericentre, respectively; the unprimed quantities refer to the ring particle and the primed quantities refer to the satellite's orbital elements. If the objects are in resonance at least one of these arguments has restricted values and is said to librate rather than circulate through 360° . The coefficients associated with each of these arguments in the expansion of the perturbing potential have terms proportional to e and e' , respectively, where e is the eccentricity. Using the terminology of Goldreich and Tremaine⁴, libration of θ_1 corresponds to an eccentric resonance if $e \neq 0$ and a Lindblad resonance if $e \approx 0$, and libration of θ_2 corresponds to a corotation resonance. It is the eccentric or Lindblad resonances that are usually associated with confinement of the ring edges^{4,6,7}. Because the precession of the pericentres, $\dot{\omega}$ and $\dot{\omega}'$, is dominated by the oblateness of Uranus, the precise semimajor axis a of each satellite resonance (the value of a such that $\dot{\theta} = 0$) and the distance between the two resonances are determined by the gravity coefficients $J_2, J_4, \dots$ for the planet, as well as by GM_U where G is the universal gravitational constant and M_U is the mass of Uranus⁶.

Porco and Goldreich⁶ showed that Cordelia is involved in three first-order eccentric resonances with ring edges: (1) a 24:25 outer resonance with the inner edge of ϵ ring, (2) a 122:123

outer resonance with the λ ring and (3) a 23:22 inner resonance with the δ ring. In addition, the proximity of the Cordelia 13:12 inner resonance to the 3:2 inner Cressida resonance and the effects on the η ring need to be re-examined using revised satellite parameters^{8,9}.

Using Cordelia as a paradigm, we calculated the locations of all first-order inner and outer eccentric resonances associated with a hypothetical satellite (HS) placed at 1-km intervals in the range $38,000 \leq a_s \leq 54,000$ km. For each value of the satellite's semimajor axis a_s , the locations of the resonances were compared with the edges of the 6, 5, 4, α , β , η , γ , δ , λ and ϵ rings as determined from occultation data^{1,2}, because these data have produced semimajor axes accurate to ≤ 1 km. The values of a_s for which a hypothetical satellite was within 1 km of a resonance with the appropriate edges (inner edge for an interior satellite and outer edge for an exterior satellite) of three or more rings are given in Table 1, together with the appropriate resonances; satellite HS22 is < 2 km from Cordelia's orbit and can be identified with that satellite to within observational accuracy. Using the torque formulae given by Goldreich and Porco⁷, we have calculated that, apart from the separate problem of atmospheric drag, small satellites ($r \leq 10$ km) at these locations can exert sufficient torques to prevent ring spreading.

FIG. 2 Binned values of I/F (where I is the absolute intensity and πF is the incident solar flux density) as a function of radius for a $\sim 3^\circ$ reduced smear section of the image shown in Fig. 1. The raw image was radiometrically and geometrically corrected and navigated using standard techniques¹⁹. The radial locations (not semimajor axes) of rings that have counterparts in backscattered images are indicated. The semimajor axes of all first-order eccentric resonances with the known satellites of Uranus are shown at the top, where solid lines are resonances with Cordelia and dashed lines are resonances with other satellites of Uranus. For clarity the large number of Cordelia resonances in the range $48,500 \leq a \leq 51,000$ km have been omitted. The large dashed lines denote the semimajor axes of the 25 hypothetical satellites discussed in the text. The darker shading denotes the radial extent of regions within the main ring system that contain little or no visible material. *a*, The region $39,600 \leq a \leq 45,800$ km. *b*, The region $45,800 \leq a \leq 52,000$ km. The location for the η ring denotes that of the broad component. Note the difference in the vertical scale between *a* and *b*. The oscillations visible in the innermost section of the profile ($39,500$ – $40,500$ km) are artefacts of our binning procedure caused by the small number of pixels in this region ($< 5\%$ of our total sample).



Although it is possible to select a number of sets of at least nine satellites which together can bound all the major rings of Uranus (for example, HS1, 5, 6, 7, 11, 15, 22, 23 and Ophelia), certain satellites have key positions^{20,21}. For example, HS1 occupies the only orbit inside ring 6 that can provide outer eccentric resonances to shepherd rings 6, 5 and 4, whereas HS6 can provide inner eccentric resonances for the same rings. The locations of two of these hypothetical satellites, HS12 at 46,950 km and HS16 at 47,921 km, lie close to the centres of cleared regions as discussed above and each of these satellites has resonances associated with four ring edges. The same procedure was used by Horn *et al.*²³ to investigate possible orbits for a moonlet thought to be producing a density wave on the δ ring. Although it was subsequently rejected on other grounds, one of their list of six possible candidates corresponds to HS16. We have recently become aware of a study by Gresh²⁴ which used a technique similar to ours and which also identified the importance of a satellite at the location of HS6: another key satellite position at 47,166 km was suggested by Gresh, but this does not meet the tolerance restrictions of our study. We believe that the combination of our results with those from these other studies provides convincing evidence for the existence of small satellites at specific locations in the rings of Uranus.

The Cordelia connection

There is a remarkable correlation between the semimajor axes of our hypothetical satellites and the locations of first-order resonances with Cordelia (Fig. 2 and Table 2). If we consider the region $38,000 \leq a \leq 48,500$ km (with the upper limit chosen to eliminate the saturation of first-order resonances as the orbit of Cordelia is approached), then Cordelia has 47 first-order eccentric or corotation resonances in this region with values of p in the range $2 \leq p \leq 25$. Five of the 20 hypothetical satellites (HS6, 11, 12, 13 and 14) in the region are within 12 km, and three (HS11, 12 and 13) within 5 km, of a first-order resonance with Cordelia; furthermore, HS6 lies in a low-contrast region relative to the background 'ramp' (see Fig. 2a). Although it is not possible for all of these satellites to be in resonance with Cordelia using currently adopted values for the relevant orbital elements, changes to the adopted value of Cordelia's semimajor axis of $\sim |a_{\text{res}} - a_s|$ would produce a number of resonant pairs.

We have used a Monte Carlo technique to estimate the probability of such a configuration occurring by chance. We generated 100,000 sets of 20 satellites, randomly located in a uniform distribution within the given range of semimajor axis; for each set we noted the number of satellites lying within specified distances of a first-order resonance with Cordelia. The mean numbers within 12 and 5 km were found to be 1.03 ± 0.99 and 0.40 ± 0.62 , respectively. This suggests that the association with Cordelia is significant at more than 4σ level in each case and that the hypothetical satellites that could bound the rings are not randomly distributed.

Furthermore, we note that Cordelia is itself very close to a resonance with Rosalind. The six separate arguments²⁵ of the

5:3 Rosalind resonance have associated semimajor axes ranging from 49,724.9 km to 49,786.7 km, whereas the semimajor axis of Cordelia is 49,752 km (ref. 9). The closest resonance to Cordelia is the 5:3 Rosalind I'^2 resonance at 49,746.5 km. The only other first- or second-order Rosalind resonances in the ring system are the 2:1 (or 4:2) resonances between 44,014.4 km and 44,119.1 km.

If the resonances involving Cordelia and known or hypothetical satellites exist then it is likely that they came about owing to orbital evolution. It is now recognized that the five major satellites of Uranus are tidally evolved²⁵. The rate of change of semimajor axis of a satellite due to the tides it raises on Uranus is given by²⁶

$$\left(\frac{da}{dt}\right)_t = \text{sign}(N - n) \frac{8.9 \times 10^{32} (m/M)}{Q} \frac{1}{a^{11/2}} \text{ km yr}^{-1} \quad (3)$$

where Q is the tidal dissipation function, N the spin frequency and M the mass of Uranus, and m , n and a are the mass, mean motion and semimajor axis (in km) of the satellite. Nine of the ten satellites discovered by Voyager 2 are inside the synchronous orbit ($n > N$) and tidal effects will cause their orbits to decay towards the planet.

The orbits of two satellites inside the synchronous orbit will approach each other provided

$$(m'/m)(a/a')^{11/2} > 1 \quad (4)$$

If this condition is satisfied then capture into resonance is possible²⁵. This is the case for the Rosalind–Cordelia pair, as well as Cordelia and any small satellite ($r < 10$ km) orbiting in the vicinity of the rings. It is also true for the Cressida–Bianca pair, which could be in a 32:30 or 16:15 resonance. Recent work by Tittlemore and Wisdom²⁷ suggests that $11,000 < Q < 39,000$ for Uranus; using equation (3) this implies a decay of ≥ 250 km in Cordelia's orbit over the age of the Solar System. But estimates of the transfer of angular momentum from Cordelia's interaction with the ϵ ring¹⁶ suggest rates at least an order of magnitude larger than those due to tidal effects, implying net motion away from Rosalind, but still towards any interior small satellites. This does not preclude capture into a 5:3 Rosalind resonance because Q could have been smaller and $(da/dt)_t$ larger in the past. In any case the claimed maximum lifetime of 6×10^8 yr for the ϵ ring¹⁶ has to be re-examined to allow for encounter with this resonance.

Although it is possible that the structure visible in the rings of Uranus could be maintained by a large number of small satellites ($r \ll 1$ km) orbiting close to the rings they shepherd, we consider it more plausible that a smaller number of larger satellites at key locations act as shepherds to three or more rings, in a similar fashion to Cordelia. We believe that the observations of gaps in the high-phase-angle image (Fig. 1) and a density wave on the δ ring²³ have already provided further indirect evidence for the existence of at least two satellites in the main ring system: (1) a satellite of radius 9 km at $a = 46,950$ km and (2) a satellite of radius 9 km at $a = 47,921$ km. $\square$

Received 31 August; accepted 9 November 1990.

- French, R. G. *et al.* *Icarus* **73**, 349–378 (1988).
- Elliot, J. L. & Nicholson, P. D. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 25–72 (University of Arizona Press, Tucson, 1984).
- Gresh, D. L. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 589–637 (University of Arizona Press, Tucson, 1984).
- Goldreich, P. & Tremaine, S. *A. Rev. Astr. Astrophys.* **20**, 249–283 (1982).
- Smith, B. A. *et al.* *Science* **233**, 43–64 (1986).
- Porco, C. C. & Goldreich, P. *Astr. J.* **93**, 724–729 (1987).
- Goldreich, P. & Porco, C. C. *Astr. J.* **93**, 730–737 (1987).
- Thomas, P., Weitz, C. & Veerka, J. *Icarus* **81**, 92–101 (1989).
- Dermott, S. F. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 1268–1271 (1987).
- Owen, W. M. & Synnott, S. P. *Astr. J.* **93**, 1268–1271 (1987).
- Cuzzi, J. & Scargle, J. D. *Astr. J.* **292**, 276–290 (1985).
- Showalter, M. *IAU Circ. No.* 5052 (1990).
- Spahn, F. & Sponholz, H. *Nature* **339**, 607–608 (1989).
- Lane, A. L. *et al.* *Science* **233**, 65–70 (1986).
- Colwell, J. E. *et al.* *Icarus* **83**, 102–125 (1990).
- Broadfoot, A. L. *et al.* *Science* **233**, 74–79 (1980).

- Esposito, L. W. & Colwell, J. E. *Nature* **339**, 605–607 (1989).
- Colwell, J. E. & Esposito, L. W. *Icarus* **86**, 530–560 (1990).
- Wisdom, J. *Astr. J.* **85**, 1122–1133 (1980).
- Murray, C. D. & Thompson, R. P. *Vistas Astr.* **32**, 225–234 (1988).
- Murray, C. D. & Thompson, R. P. in *Dynamics of Astrophysical Disks* (ed. Sellwood, J.) 17–18 (Cambridge University Press, 1989).
- Murray, C. D. & Thompson, R. P. *Bull. Am. astr. Soc.* **21**, 1015 (1989).
- Ockert, M. E., Cuzzi, J. N., Porco, C. C. & Johnson, T. V. *J. geophys. Res.* **92**, 14969–14978 (1987).
- Horn, L. J., Yanamandra-Fisher, P. A., Esposito, L. W. & Lane, A. L. *Icarus* **76**, 485–492 (1988).
- Gresh, D. L. thesis, Stanford University (1990).
- Dermott, S. F., Malhotra, R. & Murray, C. D. *Icarus* **76**, 295–334 (1988).
- Munk, W. H. & MacDonald, G. J. F. *The Rotation of the Earth* (Cambridge University Press, 1960).
- Tittlemore, W. C. & Wisdom, J. *Icarus* **85**, 394–443 (1990).

ACKNOWLEDGEMENTS. We thank S. Dermott, P. Nicholson and J. Papaloizou for discussions, and C. England and the staff of the Space and Atmospheric Physics Group at Imperial College, University of London for assistance. We also thank M. Showalter for comments and for making us aware of the work of D. Gresh. This work was supported by the SERC.

Caenorhabditis elegans *ras* gene *let-60* acts as a switch in the pathway of vulval induction

Greg J. Beitel, Scott G. Clark & H. Robert Horvitz

Howard Hughes Medical Institute, Department of Biology, Room 56-629, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

The *let-60* gene, an essential *ras* gene of the nematode *Caenorhabditis elegans*, acts as a switch in the inductive signalling pathway that initiates vulva formation. Recessive *let-60* mutations that cause a vulvaless phenotype prevent *let-60* function in response to the inductive signal. These mutations are clustered and define regions necessary either for the activation or for the action of the *let-60 ras* protein. Dominant *let-60* mutations that cause a multivulva phenotype alter codon 13 and activate *let-60 in vivo*, rendering it independent of the inductive signal. The *let-60* gene acts within an extensively defined genetic pathway, and other genes within this pathway seem likely to encode molecules that regulate *let-60* function as well as molecules that are targets of *let-60* action.

THE vulva of the *C. elegans* hermaphrodite is formed by the 22 descendants of three ectodermal blast cells, P5.p, P6.p and P7.p (Fig. 1a, b)¹. Three other cells—P3.p, P4.p and P8.p—also have the potential to generate vulval cell types. All six of these cells can undergo any of three alternative cell lineages (referred to as the 1°, 2° and 3° fates) and, because of their equivalence in developmental potential, are said to define the vulval equivalence group^{2,3}. Cells that express the 1° and 2° fates generate eight and seven vulval descendants, respectively, whereas those that express the 3° fate generate two non-vulval descendants, which fuse with the syncytial hypoderm that envelops the animal.

An inductive signal from the anchor cell in the gonad causes the nearest Pn.p cells to express 1° or 2° lineages: if no anchor cell is present, all six cells express the 3° fate, vulval development does not occur, and a vulvaless (Vul) animal is produced (Fig. 1c). Whether cells express the 1° or the 2° fate appears to be determined not only by the anchor-cell inducing signal, which is spatially graded, but also by interactions among the six cells of the vulval equivalence group³⁻⁶.

Many mutants abnormal in the vulval cell lineages have been isolated (for example, refs 7, 8). These mutants fall into two phenotypic classes: vulvaless (Vul), in which no vulva is formed, and multivulva (Muv), in which supernumerary vulva-like structures are generated. They define at least 23 genes involved in the specification of the vulval cell lineages, and many display transformations in the fates of the cells of the vulval equivalence group⁹. For example, in some Vul mutants, all six cells of the vulval equivalence group express the 3° fate; this phenotype is precisely that seen in wild-type animals lacking an anchor cell, suggesting that these mutants are defective in the inductive signalling pathway (Fig. 1c, d). By contrast, in Muv mutants all six cells can express either the 1° or 2° fates, and do so even if the anchor cell has been killed; thus, these mutants behave as if there is an inductive signal even when the inducing cell is absent (Fig. 1e, f).

Here, we describe genetic and molecular analyses of the gene

let-60 which we show to act as a switch in the pathway of vulval induction. The *let-60* gene is required for viability¹⁰. We now report the isolation of partial loss-of-function *let-60* alleles and the finding that these mutations can cause a Vul phenotype. In addition, we show that Muv mutations previously reported as defining the *lin-34* gene⁷ are dominant alleles of the *let-60* gene. Our genetic analyses reveal that *let-60* Vul mutations decrease, and *let-60* Muv mutations increase, the activity of *let-60*, suggesting that this gene is essential for vulval induction and that its product can be activated by mutation, rendering its function independent of the inductive signal.

Han and Sternberg¹¹ have recently cloned the *let-60* gene and discovered that it encodes a protein 83% identical in its first 164 amino acids to the human N-ras protein. Mutations that activate *ras* in mammalian cells are oncogenic and constitute one of the most frequent types of mutations associated with human cancers¹². The *ras* proteins, which bind GTP and are members of the G-protein superfamily¹², are thought to act as switches in signal transduction pathways in both yeast and mammalian cells^{13,14}. The function of *ras* genes in multicellular animals is less clear.

Our genetic studies reveal that the *let-60 ras* gene functions in signal transduction during inductive processes in normal development. Our molecular studies of *let-60* mutants indicate that all five dominant Muv mutants, in which this gene is activated independently of the inductive signal, contain the same point mutation, changing the glycine at codon 13 to glutamic acid. Gly 13 is one of several sites in mammalian *ras* genes that, when altered by mutation, can cause oncogenic activation^{12,15}. Seven Vul and/or lethal mutants in which *let-60* is partially inactivated are altered in regions of the *ras* protein that are not essential for an activated *ras* protein to transform cells in culture^{16,17}. These mutations might define sites necessary for normal *ras* activation but not for the function of a *ras* protein once activated. One *let-60* lethal mutation affects amino acid 37, which is within the effector domain of *ras* proteins^{16,17}; this mutant might be defective in the interaction between the *let-60 ras* protein and a target of *ras* activity.

let-60 Vul mutations prevent vulval induction

To identify new genes involved in vulval development, we isolated mutations that suppress the Muv phenotype caused by loss-of-function mutations in *lin-15*. Of 156 such suppressor mutations, seven are alleles of the gene *let-60* (Table 1 legend). The *let-60* gene was originally defined by three recessive larval lethal mutations (*s59*, *s1124*, *s1155*) isolated in screens for lethal mutations near the *unc-22* gene on chromosome IV^{10,18}. Two of the new *let-60* mutations (*n1531* and *n2031*) are dominant suppressors of the mutation *lin-15(n765)*; the other five (*n1876*, *n2021*, *n2022*, *n2034*, *n2035*) are recessive suppressors. All seven of these suppressor mutations cause at least some lethality in homozygotes (in either a *lin-15(n765)* or a *lin-15(+)* background), and only *n2021* animals are viable as a homozygous strain (Table 1a). (The recessive suppression of the *lin-15* Muv phenotype can be scored because some *let-60* homozygous progeny of *let-60/+* animals heterozygous for recessive *let-60* alleles survive; Table 1a). Many of the dying *let-60* animals arrest development as rigid, rod-like second larval stage (L2) larvae. These seven suppressor mutations fail to complement for

lethality, and the gene they define maps within 0.02 cM of the *dpy-20* gene on chromosome IV (Fig. 2). Clark *et al.*¹⁸ identified 31 essential genes in this region, including *let-60*. Our suppressor mutations *n1876* and *n1531* fail to complement the mutations *let-60(s59)* and *let-60(s1124)*, respectively, for lethality, indicating that our suppressors are alleles of the *let-60* gene.

All seven suppressor mutations (in either a *lin-15(n765)* or a *lin-15(+)* background) result in a variable mutant phenotype: some animals die as rod-like L2 larvae, other animals appear wild-type, and others are vulvaless (Vul) (Table 1a). The two dominant suppressor mutations result in a dominant Vul

phenotype (which is somewhat cold-sensitive; Tables 1a, 2a and Table 1a legend), whereas the five recessive suppressor mutations result in a recessive Vul phenotype. Hereafter we refer to these mutations as dominant Vul and recessive Vul *let-60* alleles, respectively.

Different *let-60* alleles seem to affect viability and vulva formation differentially (Table 1a). For example, 57% of *n1876/n1876* homozygotes die and the other 43% are Vul; 53% of *n2022/n2022* homozygotes die and 7% are Vul; 1% of *n2021/n2021* homozygotes die and 14% are Vul; and almost all *s59/s59* homozygotes die and the rare survivors are wild-type and not Vul. Thus, the

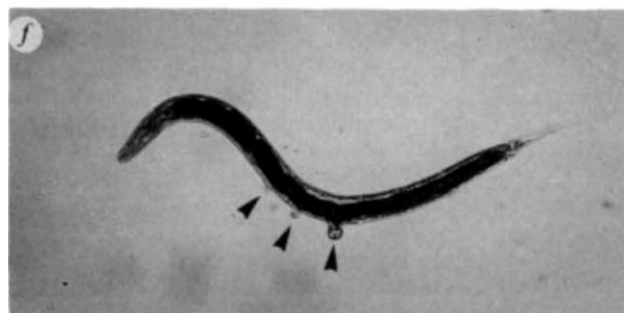
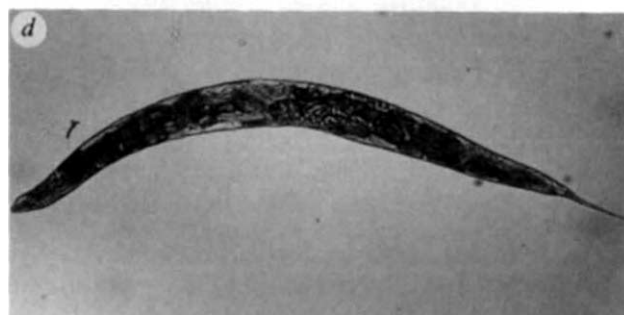
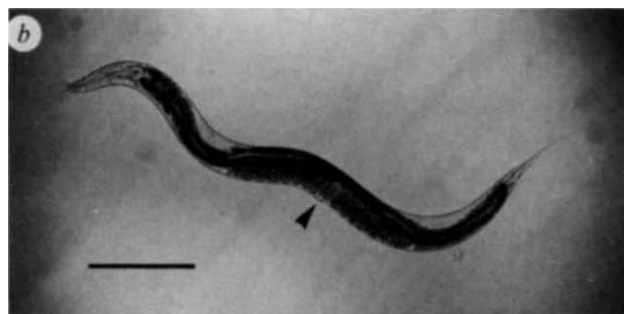
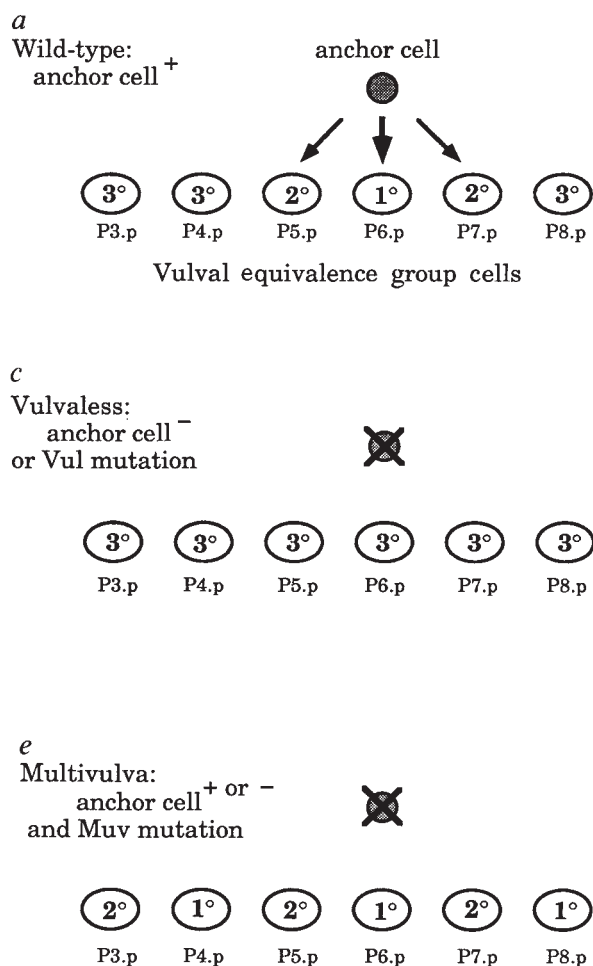


FIG. 1 Models for one aspect of vulval development in wild-type, vulvaless (Vul) and multivulva (Muv) animals, based upon refs 1–3, 9, 19. In addition to the cell interactions shown, other cell interactions are probably important in specifying vulval cell fates^{4,6}. **a**, Wild-type. A signal (or signals) from the gonadal anchor cell induces the nearest Pn.p cells to express 1° or 2° lineages, which generate vulval descendants; the 3° lineage generates non-vulval descendants. In addition, interactions among the Pn.p cells might prevent two neighbouring Pn.p cells from both expressing the 1° lineage⁴, and interactions between the syncytial hypodermis and the Pn.p cells might prevent uninduced Pn.p cells from expressing 1° and 2° lineages⁶. **b**, Bright-field photomicrograph of a wild-type adult hermaphrodite. The arrow indicates the vulva. Bar, 0.2 mm. **c**, In vulvaless (Vul) animals all six cells of the vulval equivalence group express the 3° lineage, and no vulva is formed. Wild-type animals in which the gonadal anchor cell has been killed using a laser microbeam become Vul because the vulval inducing signal has been eliminated. Mutant animals defective in the genes *lin-2*, *lin-3*, *lin-7*, *lin-10*, *let-23*, or *let-60* are Vul apparently because the inductive signalling pathway is disrupted (refs 5, 9 and Table 1b). **d**, Bright-field photomicrograph of a *let-60(n1876)* Vul adult hermaphrodite. No vulva is present, despite the presence of the gonadal anchor cell in the animal. Both hatched larvae and

eggs in a late stage of development can be seen within this animal. **e**, In multivulva (Muv) animals all six cells of the vulval equivalence group can express the 1° or 2° lineage, and extra vulval-like structures are formed. Mutant animals abnormal in a variety of genes, including *lin-15* and *let-60*, are Muv because vulval cell lineages can be activated in all six cells of the vulval equivalence group whether or not the gonadal inductive signal is present. The Muv phenotype of *lin-15* animals does not depend on the inductive signal, because if the gonad of a young *lin-15* animal is killed using a laser microbeam the mature animal still expresses the Muv phenotype⁹. We have made similar observations for the dominant Muv *let-60* alleles *n1046*, *n1700* or *n1849*. The *lin-15* product seems to act within the syncytial hypodermis and lack of *lin-15* function activates vulval cell lineages, possibly by preventing an interaction between the hypodermis and the cells of the vulval equivalence group⁶. The site of action of *let-60* has not been established. **f**, Bright-field photomicrograph of a *let-60(n1046)* Muv adult hermaphrodite in which the somatic gonadal primordium, which generates the anchor cell, was destroyed in the first larval stage using laser microsurgery²². Arrows indicate positions of the multiple vulval-like structures, which formed despite the absence of an inductive signal from an anchor cell in this animal.

TABLE 1 *let-60* Vul and lethal mutants

(a) Penetrances of <i>let-60</i> Vul and lethal phenotypes								
<i>let-60</i> allele	<i>m/+</i>		<i>m/m</i> (F1)			<i>m/m</i> (F2)		
	Vul	WT	Vul	Lethal	WT	Vul	Lethal	WT
Dominant Vul								
<i>n1531</i>	100%	0%	0%	100%	0%	NA	NA	NA
<i>n2031</i>	27%	73%	0%	100%	0%	NA	NA	NA
Revertant								
<i>n1981 n1046</i>	0%	100%	0%	100%	0%	NA	NA	NA
Recessive Vul								
<i>n1876</i>	0%	100%	43%	57%	0%	0%	100%	0%
<i>n2034</i>	0%	100%	47%	53%	0%	0%	100%	0%
<i>n2035</i>	0%	100%	51%	47%	2%	0%	100%	0%
<i>n2022</i>	0%	100%	7%	53%	40%	0%	100%	0%
<i>n2021</i>	0%	100%	14%	1%	85%	18%	52%	30%
Recessive lethal								
<i>s1155</i>	0%	100%	0%	100%	0%	NA	NA	NA
<i>s59</i>	0%	100%	0%	100%	0%	NA	NA	NA
<i>s1124</i>	0%	100%	0%	100%	0%	NA	NA	NA
(b) Cell lineages of Pn.p cells in wild-type and <i>let-60</i> Vul mutants								
	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p		
wild-type (<i>ac</i> ⁺)	3°	3°	2°	1°	2°	3°		
wild-type (<i>ac</i> ⁻)	3°	3°	3°	3°	3°	3°		
<i>let-60(n1531)/+</i>	3°	3°	3°	3°	3°	3°		
<i>let-60(n1876)</i>	3°	3°	3°	3°	3°	3°		

a. Penetrances of the Vul and lethal phenotypes of 11 *let-60* alleles. *m/+* heterozygotes between the mutant allele and the wild-type allele; *m/m*, homozygotes for the mutant allele; *m/m* (F1), homozygotes generated from *m/+* heterozygous parents; *m/m* (F2), homozygotes generated from *m/m* homozygous parents; Vul, vulvaless phenotype. Lethal, larval-arrest or sterile phenotype, animals die as rod-like early L2 larvae, but F1 homozygotes of the following genotypes sometimes grow beyond the early L2 stage: *n1981 n1046* (40% rod-like L2s, 60% L2-L4s); *s59* (20% rod-like L2s, 80% L2-L4s); *s1124* (90% rod-like L2s, 10% L2-L4s); *s1155* (5% rod-like L2s, 10% L2-L4s, 85% sterile adults of which 95% are Vul). WT, wild-type phenotype; NA, not applicable, because no *m/m* homozygotes survive. Between 200 and 650 animals were scored in each experiment. All phenotypes were scored at 20 °C unless otherwise indicated. For the dominant Vul mutations *n1531* and *n2031*, *m/+* hermaphrodites were identified as non-Dpy progeny from *m/dpy-20* parents; as all *m/m* progeny of these genotypes die as larvae, the numbers of surviving Dpy, Vul non-Dpy, and non-Vul non-Dpy progeny allowed us to estimate the number of inviable *m/+* animals based upon the 2:1 ratio of non-Dpy:Dpy animals expected if all *m/+* animals survived. Both *n1531* and *n2031* display cold-sensitivity for the lethal and Vul phenotypes: of *n1531/+* animals, at 15 °C 65% are Vul and 35% are lethal, at 20 °C 100% are Vul, and at 25 °C 44% are wild-type and 56% are Vul; of *n2031/+* animals, at 15 °C 63% are Vul and 37% are lethal, at 20 °C 73% are wild-type and 27% are Vul, and at 25 °C 99% are wild-type and 1% are Vul. For the other nine mutations listed, *m/+* hermaphrodites were identified as Unc progeny from *m/nT1 n754* parents; *nT1* is a reciprocal translocation that balances *let-60*, and the *n754* mutation is linked to this translocation and causes a dominant uncoordinated phenotype (ref. 7 and unpublished results). F1 *m/m* hermaphrodites were identified as non-Unc progeny from *m/nT1 n754* parents. F2 *m/m* hermaphrodites were obtained as progeny of F1 *m/m* hermaphrodites. The *let-60* alleles *s59*, *s1124* and *s1155* were obtained from Clark *et al.*¹⁸. We isolated the seven *let-60* mutations *n1531*, *n2031*, *n1876*, *n2034*, *n2035*, *n2022* and *n2021* as suppressors of the Muv mutation *lin-15(n765ts)* after mutagenesis with EMS³² and examination of the progeny of 38,000 F1 animals. Most of the screen was done by placing six F1 progeny of mutagenized hermaphrodites on single Petri dishes, so that suppressors inviable or genetically intractable (for example, generating completely penetrant Vul hermaphrodites that do not mate) as homozygotes could be isolated by rescuing the mutation from heterozygous siblings of suppressed animals. The mutation *n1981* was isolated as a *cis*-dominant suppressor of the Muv phenotype caused by the *let-60(n1046)* mutation. Specifically, using the *lin-8(n111)* mutation³³ as an enhancer of the Muv phenotype of the *let-60* Muv mutations *n1046*, *n1700* and *n1849* (data not shown), we mutagenized *lin-8;let-60(Muv)* strains with EMS and sought strains that when grown at 25 °C were reduced in the penetrance of the Muv phenotype. From 30,000 F1 animals we obtained five unlinked suppressors and one linked suppressor, *n1981*. The mutation *n1981* maps within 0.03 cM of *n1046*, as demonstrated by our failure to find any Muv recombinants among ~10,000 F1 progeny of animals of genotype *egl-21(n2057) n1981 n1046 unc-22(e66)/dpy-20(e1282)*. (The *egl-21* mutation was induced in the same experiment that generated the *n1981* mutation). b. Vulval cell lineages of wild-type and *let-60* Vul mutants. The lineages of the six cells of the vulval equivalence group, P3.p-P8.p, were described in intact wild-type animals (*ac*⁺)¹ and in wild-type animals lacking the gonadal anchor cell (*ac*⁻)³ (Fig. 1 legend). 1°, 2° and 3° indicate lineages that generate eight, seven or two descendants, respectively; other criteria for distinguishing these lineages are described by Sternberg and Horvitz³. The 1° and 2° lineages generate vulval cells, and the 3° lineages generate non-vulval cells. Nine animals of each genotype were scored. All cells expressed the lineages shown, except for the P6.p cells in two *n1531/+* animals, which divided multiple times expressing an apparent 1° lineage; such incomplete penetrance is typical of Vul mutants^{5,9}. Cell lineages were determined by direct observation using Nomarski optics, as previously described¹.

phenotypes caused by the recessive *let-60* mutations do not correspond to a simple allelic series, and some of them might define sites within the *let-60* gene product that function with relative cell-type specificity.

Vulval cell lineages of a dominant Vul and a recessive Vul *let-60* mutant are shown in Table 1b. In both cases, the three cells P5.p, P6.p and P7.p, which are normally induced to express vulval cell lineages (1°, 2°), instead express the non-vulval cell lineage (3°) characteristic of the cells P3.p, P4.p and P8.p (Fig. 1c). Thus, these *let-60* mutants display a transformation in the vulval cell lineages identical to that seen when the inductive signal from the anchor cell is eliminated¹⁹.

Mutations in *let-23* and at least four other genes also cause both a Vul and a rod-like larval lethal phenotype (ref. 7; S.G.C.

and H.R.H., unpublished results). These observations suggest that these genes act in a pathway that functions not only during vulval development but also during an earlier phase of larval development. In addition, both *let-23* mutations^{7,20} and the two dominant Vul *let-60* mutations affect another cell fate that is specified by cell interactions: these mutations transform the ectoblast cell P12 to express the cell lineage normally expressed by P11 as well as affect the ability of males to mate (data not shown).

let-60 Muv mutations activate vulval lineages

The small region of chromosome IV that contains the *let-60* mutations described above also contains another class of mutation that affects the vulval cell lineages. The original mutation

TABLE 2 Penetrance of *let-60* phenotypes

(a) Gene dosage studies with dominant Muv and Vul <i>let-60</i> alleles					
	<i>n1046</i> % Muv at 20 °C	<i>n1531</i> % Vul at 20 °C	<i>n1531</i> % Vul at 25 °C	<i>n2031</i> % Vul at 20 °C	
<i>m/sDf8</i>	7% (<i>n</i> =495)	Lethal	Lethal	Lethal	
<i>m/+</i>	23% (<i>n</i> =407)	100% (<i>n</i> =276)	56% (<i>n</i> =249)	27% (<i>n</i> =211)	
<i>m/+/+</i>	53% (<i>n</i> =161)	11.1% (<i>n</i> =81)	ND	0% (<i>n</i> =120)	
<i>m/m</i>	93% (<i>n</i> =1,559)	Lethal	Lethal	Lethal	
<i>m/m/+</i>	100% (<i>n</i> =309)	100% (<i>n</i> =223)	100% (<i>n</i> =110)	100% (<i>n</i> =111)	
(b) The <i>let-60</i> dominant Muv phenotype is dominant to other <i>let-60</i> phenotypes					
	Dominant Vul <i>n1531</i>	Recessive Vul <i>n1876</i>	% Muv Recessive Let <i>s1124</i>	Wild-type +	Deficiency <i>sDf8</i>
<i>m/n1046</i>	19% (<i>n</i> =294)	14% (<i>n</i> =326)	26% (<i>n</i> =430)	12% (<i>n</i> =304)	7% (<i>n</i> =495)

a. Penetrances of the Muv, Vul and lethal phenotypes of strains containing different numbers of *let-60* wild-type or dominant mutant alleles. *m*, the *let-60* mutation *n1046*, *n1531* or *n2031*; *sDf8*, the smallest existing deficiency that deletes the *let-60* gene; +, the wild-type *let-60* allele; *n*, number of animals scored. ND, not determined. A third copy of the *let-60* locus was provided by the free duplication *nDp5*, isolated by S. Kim (personal communication). *nDp5* duplicates a region of chromosome IV that includes the genes *unc-44* and *unc-26* and spans the entire region shown in Fig. 2. Animals of *let-60* genotype *+/+/+* are phenotypically wild-type. For purposes of comparison, some of the data presented in this table are also shown in Tables 1 and 3. Animals of the genotypes indicated were constructed using linked markers and standard methods for *C. elegans* genetics³². b. Penetrances of the Muv phenotype of strains heterozygous for the dominant Muv mutation *let-60(n1046)* and the *let-60* allele indicated. *m*, the *let-60* alleles *n1531* (dominant Vul), *n1876* (recessive Vul), *s1124* (recessive lethal), + (wild-type), or *sDf8* (a deficiency). *n*, number of animals scored. The penetrance of the Muv phenotype of *let-60* (dominant Muv)/+ animals has varied from 5%–23% in different experiments; we have not attempted to interpret differences that fall close to or within this range.

of this class, *n1046*, was identified in a general screen for Muv and Vul mutants⁷. The dominant Muv mutation *n1046* defined a locus previously called *lin-34*. More recently, four mutations very tightly linked to *n1046* (within 0.02 cM; Fig. 2 legend), that confer a dominant Muv phenotype similar to that of *n1046* animals, have been isolated by us (Fig. 2 legend) and by others (ref. 21; and G. Jongeward and P. Sternberg, personal communication): *n1700*, *n1849*, *sy103* and *sy130*. These five closely linked dominant Muv mutations behave similarly in genetic experiments, for example, as homozygotes, as heterozygotes with each other, as heterozygotes in *trans* to a wild-type allele, as heterozygotes in *trans* to deficiencies, or in the presence of the amber suppressor *sup-7* (Table 3). These observations suggest that all five mutations are alleles of the same gene.

The *n1046* Muv mutation causes the three cells P3.p, P4.p and P8.p, which normally do not express vulval cell lineages, to express vulval cell lineages⁹. The other four closely linked Muv mutations result in Muv phenotypes that are indistinguishable at the level of the dissecting microscope. Using a laser microbeam²², we destroyed the somatic gonad in L1 animals of

genotypes *n1046*, *n1700* and *n1849* (number of animals = 13, 8 and 4, respectively) and found that in 24 of these 25 animals the Muv phenotype was still expressed. Thus, these mutations activate vulval cell lineages independently of the gonadal inductive signal (Fig. 1f).

These dominant Muv mutations do not cause a loss of gene function, as heterozygotes for any of the nine deficiencies that delete the region in which they are located (ref. 10, 18; also D. Clark and D. Baillie, personal communication) do not display the Muv phenotype. To determine the phenotype caused by loss-of-function mutations in the gene defined by these Muv mutants, we sought mutations within this gene that reduced or eliminated the gene activity responsible for the Muv phenotype. We mutagenized homozygous *n1046* animals and isolated animals in which the Muv phenotype was suppressed (Table 1a). We obtained one apparently intragenic suppressor mutation, *n1981*, which is tightly linked to *n1046* (Table 1a) and does not result in a dominant Muv phenotype but rather causes a recessive rod-like larval lethal phenotype, like that conferred by mutations in *let-60*. The double mutation *n1981 n1046* fails

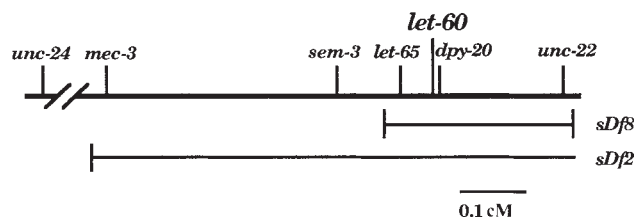


FIG. 2 Genetic map of the *let-60* IV region and map data. The *let-60* mutations that result in an early larval lethality were shown to fail to complement the deficiencies *sDf2* and *sDf8* and to map to the right of *let-65* and very close to *dpy-20* on chromosome IV¹⁸. Our data show that one of our new recessive *let-60* alleles, *n1876*, isolated as a suppressor of the Muv phenotype of the mutation *lin-15(n765)* maps within ~0.02 cM of *dpy-20*; also, the Muv mutations *n1046*, *n1700* and *n1849* map ~0.02 cM left of *dpy-20*. The Muv mutation *sy103* maps close to and left of *dpy-20*, the Muv mutation *sy130* maps close to *dpy-20*, and other recessive *let-60* alleles map between *let-65* and *dpy-20* (ref. 21). The Muv mutation *n1046* is described in ref. 7. The other Muv mutations were isolated as suppressors of various Vul mutations: *n1700* as a *lin-10* suppressor (D. Parry, S. K. Kim and H.R.H., unpublished results), *n1849* as a *let-341* suppressor (S.G.C. and H.R.H., unpublished results), *sy103* as a *let-23* suppressor (G. Jongeward and P. Sternberg, personal communication), and *sy130* as a *trans*-dominant suppressor of *let-60(sy94)* (a dominant Vul mutation)²¹. Three-factor

Three-factor Crosses				
Gene	Genotype of heterozygote	Phenotype of selected recombinants	Genotype of selected recombinants (with respect to unselected marker)	
<i>let-60(n1876)</i>	+ <i>n1876</i> + <i>sem-3</i> + <i>dpy-20</i>	Dpy	0/42	<i>n1876</i> /+
	+ <i>n1876</i> + <i>let-65</i> + <i>unc-22</i>	Unc	5/7	<i>n1876</i> /+
	<i>n1876</i> +/+ + <i>dpy-20</i> <i>unc-22</i>	Unc	15/15	<i>n1876</i> /+
<i>let-60(n1046)</i>	+ <i>n1046</i> + <i>sem-3</i> + <i>dpy-20</i>	Dpy	1/12	<i>n1046</i> /+
	+ <i>n1046</i> + <i>let-65</i> + <i>dpy-20</i>	Dpy	0/2	<i>n1046</i> /+
	+ <i>n1046</i> + <i>let-65</i> + <i>unc-22</i>	Unc	11/13	<i>n1046</i> /+
	++ <i>n1046</i> + <i>unc-24 mec-3</i> + <i>dpy-20</i>	Dpy	0/22	<i>n1046</i> /+
<i>let-60(n1700)</i>	++ <i>n1700</i> + <i>unc-24 mec-3</i> + <i>dpy-20</i>	Dpy	1/22	<i>n1700</i> /+
	++ <i>n1700</i> + <i>let-65</i> + <i>unc-22</i>	Mec	17/18	<i>n1700</i> /+
	++ <i>n1700</i> + <i>let-65</i> + <i>unc-22</i>	Unc	8/12	<i>n1700</i> /+
<i>let-60(n1849)</i>	++ <i>n1849</i> + <i>unc-24 mec-3</i> + <i>dpy-20</i>	Dpy	1/26	<i>n1849</i> /+
	++ <i>n1849</i> + <i>let-65</i> + <i>unc-22</i>	Mec	4/4	<i>n1849</i> /+
	++ <i>n1849</i> + <i>let-65</i> + <i>unc-22</i>	Unc	7/13	<i>n1849</i> /+

crosses were performed as described³². From heterozygotes of genotype *ab/c*, recombinants A non-B and B non-A were picked. The progeny of each recombinant hermaphrodite were examined for the expression of the *trans* marker *c*.

TABLE 3 The *let-60* Muv alleles are genetically similar

	<i>n1046</i>	<i>n1700</i>	<i>n1849</i>	<i>sy103</i>	<i>sy130*</i>
<i>m/m</i>	93% (<i>n</i> =1,559)	96% (<i>n</i> =703)	94% (<i>n</i> =601)	85% (<i>n</i> =732)	98% (<i>n</i> =827)
<i>m/+</i>	23% (<i>n</i> =407)	14% (<i>n</i> =442)	9% (<i>n</i> =627)	8% (<i>n</i> =267)	16% (<i>n</i> =347)
<i>m/sDf2</i>	0.5% (<i>n</i> =197)	0.8% (<i>n</i> =255)	0.6% (<i>n</i> =869)	0.4% (<i>n</i> =268)	0.0% (<i>n</i> =317)
<i>m/sDf8</i>	7% (<i>n</i> =495)	7% (<i>n</i> =207)	8% (<i>n</i> =287)	5% (<i>n</i> =215)	3% (<i>n</i> =188)
<i>m/n1046</i>	93% (<i>n</i> =1,559)	95% (<i>n</i> =258)	91% (<i>n</i> =261)	85% (<i>n</i> =209)	89% (<i>n</i> =239)
<i>m; sup-7</i>	11% (<i>n</i> =284)	29% (<i>n</i> =371)	26% (<i>n</i> =296)	34% (<i>n</i> =272)	69%* (<i>n</i> =297)

Penetrances of the Muv phenotype of the five dominant Muv *let-60* alleles. All phenotypes were scored at 20 °C except for those of strains containing the amber suppressor *sup-7(st5)* (ref. 34), which were scored at 22.5 °C. The penetrances of the Muv phenotype of *m/m; sup-7(+)* animals are indistinguishable at 22.5 °C and 20 °C. *m*, the *let-60* dominant Muv allele *n1046*, *n1700*, *n1849*, *sy103* or *sy130*. *n*, number of animals scored. The precise magnitude of the Muv penetrance varies from experiment to experiment, and we do not believe that any of the variation seen among the five alleles is meaningful. All five homozygous *m/m* strains display a slight temperature-sensitivity for the Muv phenotype (data not shown). The differences between the penetrances of the Muv phenotypes in heterozygotes for the two deficiencies *sDf2* and *sDf8* (ref. 35) are reproducible and might be a consequence of the different extents of these deficiencies (Fig. 2). The effects of the smaller deficiency *sDf8* are likely to reflect the effects of eliminating *let-60* gene activity, and indeed in various heterozygous combinations the effects of strong reduction-of-function alleles of *let-60* are more similar to those of *sDf8* than to those of *sDf2* (compare with Tables 2 and 3). Animals of the genotypes indicated were constructed using standard methods for *C. elegans* genetics³². Strains were confirmed to be homozygous for *sup-7(st5)* by showing they are sterile at 15 °C³⁶ and, in some cases, by crossing with animals carrying the *unc-93(e1500 n234)* amber mutation³⁷.

* All data for *sy130* were obtained using strains that contained the closely linked mutation *dpy-20(e1282)*, which was present on the chromosome on which *sy130* was generated²¹ and from which *sy130* has not been separated; the relatively low level of suppression seen in the *sy130; sup-7(st5)* strain is probably caused by the presence of the *dpy-20* mutation, as the penetrance of the Muv phenotype in a strain of genotype *n1700 dpy-20(e1282); sup-7(st5)* is 65% (*n*=684). As these Muv mutations are not amber mutations (Fig. 4a), we suspect the suppression caused by the *sup-7(st5)* mutation is indirect and similar to the ability of a variety of mutations, including *dpy-20(e1282)*, *unc-22(e66)* and *lon-2(e678)*, in apparently unrelated genes to affect the penetrance of the *let-60* Muv phenotype (G.J.B. and H.R.H., unpublished results).

to complement mutations in *let-60* for this lethal phenotype, establishing that *n1981* is an allele of *let-60*. These observations suggest that *n1046*, and presumably the other four closely linked dominant Muv mutations as well, are alleles of *let-60*. Our molecular data confirm this interpretation.

let-60 acts as a binary switch

The dominant Vul and the dominant Muv *let-60* mutations result in opposite transformations in vulval cell fates. The wild-type phenotype of deficiency heterozygotes indicates that neither class of dominant *let-60* mutations simply causes a loss of *let-60* gene function. To determine how these and the other *let-60* alleles affect *let-60* gene activity we used gene dosage experiments: the phenotype conferred by a mutation that elevates an essentially wild-type gene activity is often enhanced by an additional wild-type allele; the phenotype conferred by a mutation that antagonizes wild-type gene activity is often diminished by an additional wild-type allele; and the phenotype conferred by a mutation that decreases gene activity is often enhanced if the level of mutant gene activity is diminished, as in animals heterozygous for the mutation and a deletion.

We constructed strains carrying different copy numbers of wild-type and mutant *let-60* alleles (Table 2a). Increased gene dosage of either the wild-type or a Muv allele increases the penetrance of the Muv phenotype, consistent with the hypothesis that the dominant Muv mutations cause an elevation of an essentially wild-type *let-60* gene activity. Increased gene dosage of the wild-type allele decreases the penetrance of the Vul phenotype and rescues the lethal phenotype caused by dominant Vul mutations, indicating that these mutations act to antagonize wild-type *let-60* activity.

The mutant phenotypes caused by the recessive Vul alleles are enhanced when such alleles are in *trans* to a deficiency spanning *let-60*, such as *nDf27* (*nDf27* has been described⁷). For example, Vul animals homozygous for the recessive Vul allele *n1876* produce an average of 40 inviable zygotes, whereas most Vul animals of genotype *n1876/nDf27* are sterile and the rest produce an average of only five inviable zygotes (data not shown). These observations indicate that recessive Vul *let-60* alleles reduce but probably do not eliminate *let-60* activity. Of the recessive lethal *let-60* alleles, the strongest phenotype is conferred by *s1124*, which is the best candidate for totally eliminating *let-60* function (Table 1a legend).

We also examined mutant animals heterozygous for the

dominant Muv allele *let-60(n1046)* and a representative allele of each other class of *let-60* mutation—a dominant Vul, a recessive Vul, a recessive lethal, a wild-type or a deficiency allele (Table 2b). Interestingly, the dominant Muv mutation, which increases *let-60* activity, is not antagonized by a dominant Vul mutation, although the dominant Vul mutation does antagonize a wild-type allele. This observation suggests that the dominant Muv *let-60* protein fails to interact with, outcompetes or bypasses the effects of the dominant Vul *let-60* protein.

To summarize, the opposite phenotypic effects of the *let-60* Muv and Vul mutations reflect the opposite effects of these mutations on *let-60* gene activity: a reduction of *let-60* gene function prevents vulval induction, and an increase in *let-60* gene function causes activation of the vulval cell lineages even in the absence of the inducing signal. These observations indicate that *let-60* acts as a binary switch in the signalling pathway responsible for vulval induction (Fig. 3). Han *et al.*²¹ have independently reached similar conclusions.

Muv mutations are codon 13 substitutions

We determined the DNA sequences of the *let-60* coding regions of mutant animals. Using the *let-60* gene sequence from Han and Sternberg¹¹, we amplified genomic DNA from *let-60* mutants using the asymmetric polymerase chain reaction (PCR) and determined the nucleotide sequence of the amplified DNA (Fig. 4a).

The five *let-60* dominant Muv mutations are identical: in each

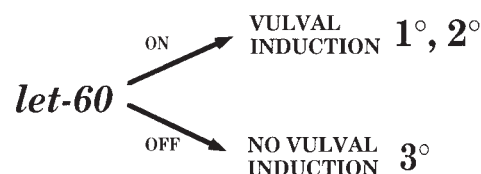


FIG. 3 Model for the action of the *let-60* gene in vulval induction. In normal development, the inductive signal from the gonad activates the *let-60* gene causing the cells P5.p, P6.p and P7.p to express the 1° or 2° vulval cell lineages; P3.p, P4.p and P8.p do not receive this signal so that *let-60* is not activated, causing these cells to express the 3° non-vulval cell lineage. In *let-60* Vul mutants, *let-60* is not activated even when the inducing signal is present. In *let-60* Muv mutants, *let-60* is activated even when no inducing signal is present.

of these mutants, a G→A transition causes a substitution of glutamic acid for glycine at codon 13. This observation establishes that these dominant Muv mutations are alleles of the *let-60* gene. The molecular identity of these five independently generated mutations explains their similarity in the genetic experiments already described. The glycine at codon 13 is one of several amino acids in mammalian *ras* proteins that when altered by point mutation can cause oncogenic activation^{12,15}. This result supports the conclusion that the Muv phenotype is a consequence of elevated *let-60* activity.

By analogy with mammalian and yeast *ras* genes, it seems likely that a variety of mutations would activate the *let-60* *ras* protein. Why then are all five of the *let-60* Muv mutations the identical base pair change? One possibility is mutagen specificity: ethyl methanesulphonate (EMS) was used as a mutagen for the generation of all five *let-60* Muv mutations, and mutagens are known to induce mutations at different sites at vastly different frequencies (refs 23, 24). For example, in mammalian *ras* genes, different mutagens induce different activating mutations: all 36 Ha-*ras*-1 mutations in rat mammary carcinomas induced by *N*-nitroso-*N*-methylurea contain the identical codon 12 G→A transition mutation, despite the fact that another mutagen can induce similar tumours through alter-

ations at a different position of the *ras* gene²⁵. In addition, other EMS-induced mutations might activate *let-60* *ras*, but perhaps only mutations at codon 13 result in a viable animal with a Muv phenotype; other *let-60* activation mutations might not affect vulva development or result in lethality.

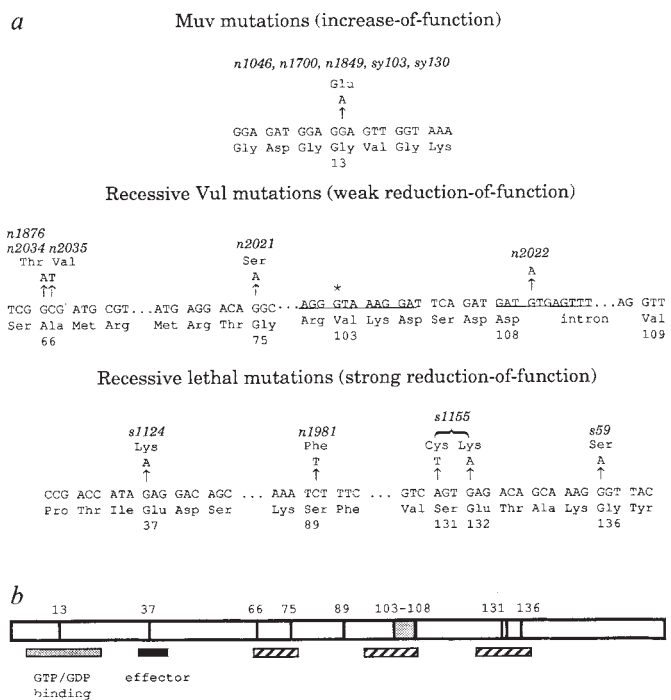
Analysis of *let-60* recessive mutations

We have also determined the DNA sequences of the *let-60* coding regions in the recessive mutants. Three of the recessive Vul mutations affect amino acid 66 of the *let-60* protein (Fig. 4a). The mutations *n1876* and *n2034* are the identical G→A transition and cause the substitution of threonine for alanine at codon 66. The mutation *n2035* is a G→A transition at the adjacent base pair, causing a different substitution, valine for alanine, at codon 66. In *n2021*, a G→A transition causes the substitution of serine for glycine at codon 75.

The effect of the fifth recessive Vul mutation, *n2022*, is less clear. This mutation is a G→A transition in the first base pair (bp) of the second intron; this G is invariant among *C. elegans* and other eukaryotic 5' splice sites (ref. 26; and C. Fields, personal communication). Although this mutation presumably inactivates the normal splice site, it does not eliminate *let-60* activity. We speculate that a cryptic splice site 18 bp 5' of the

FIG. 4a, DNA sequences of 14 *let-60* *ras* alleles. The wild-type DNA and presumptive protein sequences are from Han and Sternberg¹¹ and have been confirmed by us. As these authors have reported, the *let-60* protein is very similar to Ha-*ras* and N-*ras* proteins from its N terminus to amino acid 164; but we have noted (unpublished observations) that at its C terminus (amino acids 165–184) the *let-60* protein is highly basic and is more similar to *ras* proteins of the Ki-*ras*-2B type³⁸. Each allele name is indicated above the site of the DNA alteration detected in a strain containing that allele. Each base change and its associated amino-acid change are shown below the allele name. Codon numbers are indicated below the corresponding amino acids. The mutations are shown in three groups, as defined by our genetic studies. Muv mutations increase *let-60* gene activity, recessive Vul mutations decrease *let-60* gene activity, and recessive lethal mutations seem to reduce *let-60* gene activity further. The *n2022* mutation affects the first base pair of the second intron; based on our analysis (data not shown) using the Genetics Computer Group, Inc. (Madison Wisconsin) FIT-CONSENSUS program and *C. elegans* consensus 5' splice sites for short introns (C. Fields, personal communication), we propose that this mutation inactivates the normal donor splice site and that a cryptic site (*) 18 bp 5' to the normal site is used instead. The postulated and actual splice sites are underlined. The *n1981* mutation was isolated as a *cis*-dominant suppressor of the *n1046* mutation, and the DNA sequence of the suppressed strain revealed the presence of both the original *n1046* mutation and the mutation shown as *n1981*. The strain carrying the *s1155* allele has two mutations affecting adjacent amino acids. b, Schematic drawing of the sites of *let-60* mutations and of structural and functional regions of *ras* proteins as described by Barbacid¹². Amino acids 5–22 define one of a number of regions involved in GDP/GTP binding. Amino acids 30–42 appear to be involved in *ras* effector function. Amino acids 64–76, 93–108 and 124–138 are all in regions unnecessary for activated v-Ha-*ras* to transform NIH 3T3 cells.

METHODS. Exons of mutant genomic DNA were amplified by asymmetric PCR³⁹, using primers that flank the open reading frames (ORFs) of the four exons by 97–195 bp. The specific primers used are as follows: exon 1, 5'CCTGCTAGTACTCTCTTATCG and 5'CAATTAACAGCTCGATGTGTTTC; exons 2 and 3, 5'CCTGCTCGTGTGCTGTGAGGAGT and 5'CCGGTGTGCTAT-TTTTGGCGGCG; exon 4, 5'CTCACCAGCATGCACAAATCTGA and 5'GACGAG-CAGACGAGAGAACAAGTT. For the Muv mutations, 400 ng genomic DNA was used as template in each 100-μl reaction. Conditions for the asymmetric PCR reaction were: annealing at 61 °C for 2 min, extension at 72 °C for 1 min, denaturation at 92 °C for 1 min, all for 30–35 cycles. Each 100-μl reaction contained 10 mM Tris-HCl (pH 8.3) 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin, 200 μM each ATP, CTP, GTP, TTP, first primer 1.0 μM, limiting primer 0.05 μM, 5 units AmpliTaq polymerase (Perkin-Elmer Cetus, Norwalk Connecticut). For the recessive Vul and the recessive lethal mutations, which cannot be maintained in homozygous strains, two PCR reactions were done. The first was a normal symmetric reaction using template DNA derived either from two gravid adult hermaphrodites homozygous for the mutation (generated as the F1 progeny of heterozygotes for recessive Vul mutations) or from 20 dead larvae (for recessive lethal mutations) that had been



digested in PCR lysis buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.2, 2.5 mM MgCl₂, 0.45% NP-40, 0.45% Tween-20, 0.1 mg ml⁻¹ gelatin and 60 μg ml⁻¹ proteinase K; R. Barstead and R. Waterston, personal communication) at 60 °C for 1 h, and heated to 95 °C for 15 min before the addition of the PCR reagents. The products of this reaction were purified by agarose gel electrophoresis and the desired band cut out of a low-melting-point gel. 5 μl (~25 ng) of material from this gel was then used in an asymmetric PCR reaction as above, except that only 25 cycles were run. For sequencing, the products of two 100-μl asymmetric PCR reactions were pooled and purified using Qiagen (Studio City, California) PCR purification kits and the DNA sequences were determined using primers internal to the PCR products and Sequenase (USB, Cleveland, Ohio), essentially according to the instructions provided by the manufacturer. The primers used for sequencing each exon were: exon 1, 5'CCATCTCTATTTTACATTTTCCA, 5'CAATTTGCTCCAC-CACATGTGTG; exon 2, 5'TTCAGTTTTAAAAATTGCTCCGAG, 5'CCCTACC-AAGACCATAGGAACC; exon 3, 5'GTACCTTAATTGTACATTAATAAACA, 5'GGG-TAAAGGATTCAGATGATGTG; exon 4, 5'GACACCTAGATTGATATCCG, 5'GG-GTGGGAGTTGACACGACTG. The DNA sequences were determined of both strands of the ORFs of all four exons and of at least 10 bp beyond the 5' and 3' ends of each ORF for all 14 mutations.

normal site is used in the *let-60(n2022)* mutant, which would result in an in-frame deletion of six amino acids (codons 103–108) (see Fig. 4 legend).

Two of the recessive lethal *let-60* mutations affect a small region of the *let-60 ras* protein. The mutation *s59* is a G→A transition that causes the substitution of serine for glycine at codon 136. The *s1155* mutant strain contains two mutations, an A→T transversion that causes the substitution of cysteine for serine at codon 131 and a G→A transition that causes the substitution of lysine for glutamic acid at codon 132; we do not know whether one or both of these DNA changes is responsible for the phenotype of *s1155* strains.

The recessive lethal *let-60* mutation *s1124* is a G→A transition that causes the substitution of lysine for glutamic acid at codon 37. The revertant *let-60(n1981 n1046)* strain contains both the original activating *n1046* mutation affecting codon 13 and a second G→A mutation causing the replacement of serine with phenylalanine at codon 89, the latter of which is presumably responsible for the recessive lethal phenotype of this strain. Although *n1981* does not cause a dominant Vul phenotype, this allele is like the dominant Vul alleles in that it causes a dominant suppression of the Muv phenotype caused by the *lin-15(n765)* mutation. The dominant negative dominant Vul mutations will be analysed later.

Regions possibly involved in normal activation

The recessive *let-60* mutations affect codons 37, 66, 75, 103–108, 131 and/or 132 and 136. All of these codons except 37 are in regions of the *ras* protein that are highly conserved but seem to be unnecessary for the transforming activity of activated v-Ha-*ras* in NIH 3T3 cells^{16,17} (Fig. 4b). Deletions of amino acids 64–72, 102–108 and 131–138 result in apparently normal transforming activities in this assay, and deletion of amino acids 69–76 reduces transforming activity only slightly. Furthermore, point mutations of v-Ha-*ras* identical to the *let-60 ras* mutations *n1876* (Ala 66→Thr 66), *n2034* (Ala 66→Thr 66) and *n2021* (Gly 75→Ser 75) do not prevent the activated *ras* protein from transforming NIH 3T3 cells¹⁷. Thus, the amino acids we have found to be important for the biological activity of *let-60 ras* in *C. elegans* development are dispensable for the activity of a *ras* protein that is mutationally activated. These observations suggest that the amino acids affected in the *let-60* mutants might be important for activation but not for activity of the *ras* protein. Alternatively, mutation of these sites might decrease *ras* activity to a level insufficient for *let-60* function in *C. elegans* vulval induction and/or for viability, but sufficient for *ras*

function in the NIH 3T3 cell transformation assay.

One way in which the regions of *ras* defined by the recessive *let-60* mutations might function is by directly or indirectly influencing the interaction between *let-60 ras* and an activator protein. Several possible *ras* activator proteins have been described, including the yeast CDC25 protein²⁷, and the REP²⁸, Ras-GRF²⁹ and SCD25³⁰ proteins. The amino acids affected by the recessive *let-60* mutations (codons 66, 75, 103–108, 131–136) are candidates for being directly involved in interactions between the *let-60 ras* protein and an activator protein, as X-ray crystallographic studies of *ras* protein structure reveal all of these regions to be on the surface of the protein³¹. Alternatively, these regions of *ras* proteins might be important for adopting or maintaining an active conformation.

The codon 37 recessive lethal mutation *s1124* affects a region of the *ras* protein essential for the transforming activity of activated v-Ha-*ras*^{16,17} (Fig. 4b). This region, known as the effector domain, seems to be necessary for *ras* protein function and has been postulated to be a site of interaction with a target protein^{16,17}.

Discussion

Mutations that reduce *let-60* function prevent vulval induction, whereas mutations that activate *let-60* cause vulval development to occur independently of the inductive signal. Thus, *let-60* acts as a binary switch in the pathway of vulval induction. These findings provide evidence that *ras* genes can normally function in intercellular communication, possibly in response to extracellular signals. That the *let-60* gene is essential for viability reveals that its function is also important for early development to proceed normally. All five mutations that activate *let-60* alter amino acid Gly 13, a site that if altered can activate mammalian *ras* proteins¹⁵. Some recessive *let-60* mutations might define regions of the *let-60* protein involved in cell-type-specific activity. All but one of the recessive *let-60* mutations might define regions of the *ras* protein needed for normal activation but not for the activity of a mutationally-activated *ras* protein. These regions might be the direct targets of one or more proteins that activate *ras*, or they might be structurally important for a *ras* protein to adopt or maintain its active conformation. One recessive lethal *let-60* mutation affects a region of the *ras* protein believed to interact with a target of *ras* activity. Further study of the *let-60 ras* gene and of other genes involved in *C. elegans* vulval induction should reveal how a *ras* gene functions in normal development and should identify molecules that regulate *ras* activity as well as those that are targets of *ras* action. □

Received 17 October; accepted 8 November 1990.

1. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110–156 (1977).
2. Sulston, J. E. & White, J. G. *Dev. Biol.* **78**, 577–597 (1980).
3. Sternberg, P. W. & Horvitz, H. R. *Cell* **44**, 761–772 (1986).
4. Sternberg, P. W. *Nature* **335**, 551–554 (1988).
5. Sternberg, P. W. & Horvitz, H. R. *Cell* **58**, 679–693 (1989).
6. Herman, R. K. & Hedgecock, E. M. *Nature* **348**, 169–171 (1990).
7. Ferguson, E. L. & Horvitz, H. R. *Genetics* **110**, 17–72 (1985).
8. Ferguson, E. L. & Horvitz, H. R. *Genetics* **123**, 109–121 (1989).
9. Ferguson, E. L., Sternberg, P. W. & Horvitz, H. R. *Nature* **326**, 259–267 (1987).
10. Rogalski, T. M., Moerman, D. G. & Baillie, D. L. *Genetics* **102**, 725–736 (1982).
11. Han, M. & Sternberg, P. W. *Cell* (in the press).
12. Barbacid, M. *A. Rev. Biochem.* **56**, 779–827 (1987).
13. Toda, T. *et al. Cell* **40**, 27–36 (1985).
14. Downward, J., Graves, J. D., Warne, P. H., Rayter, S. & Cantrell, D. A. *Nature* **346**, 719–723 (1990).
15. Bos, J. L. *et al. Nature* **315**, 726–730 (1985).
16. Willumsen, B. M. *et al. Molec. cell. Biol.* **6**, 2646–2654 (1986).
17. Sigal, I. S., Gibbs, J. B., Dalozzo, J. S. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4725–4729 (1986).
18. Clark, D. V., Rogalski, T. M., Donati, L. M. & Baillie, D. L. *Genetics* **119**, 345–353 (1988).
19. Kimble, J. *Dev. Biol.* **87**, 286–300 (1981).
20. Sternberg, P. W. *Trends Neurosci.* **11**, 259–264 (1988).
21. Han, M., Aroian, R. & Sternberg, P. W. *Genetics* (in the press).
22. Avery, L. & Horvitz, H. R. *Cell* **51**, 1071–1078 (1987).
23. Prakash, L. & Sherman, F. *J. molec. Biol.* **79**, 65–82 (1973).
24. Coulondre, C. & Miller, J. H. *J. molec. Biol.* **117**, 525–567 (1977).

25. Zarbl, H., Sukumar, S., Arthur, A. V., Martin, Z. D. & Barbacid, M. *Nature* **315**, 382–385 (1985).
26. Mount, S. M. *Nucleic Acids Res.* **10**, 459–472 (1982).
27. Powers, S., O'Neill, K. & Wigler, M. *Molec. cell. Biol.* **9**, 390–395 (1989).
28. Downward, J., Riehl, R., Wu, L. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5998–6002 (1990).
29. Wolfman, Z. & Macara, I. G. *Science* **248**, 67–69 (1990).
30. Crechet, J. B. *et al. Science* **248**, 866–868 (1990).
31. Milburn, M. V. *et al. Science* **247**, 939–945 (1990).
32. Brenner, S. *Genetics* **77**, 71–94 (1974).
33. Horvitz, H. R. & Sulston, J. E. *Genetics* **96**, 435–454 (1980).
34. Bolten, S. L., Powell, A. P., Fischhoff, D. A. & Waterston, R. H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6784–6788 (1984).
35. Moerman, D. G. & Baillie, D. L. *Mutat. Res.* **80**, 273–279 (1981).
36. Waterston, R. H. *Genetics* **97**, 307–325 (1981).
37. Greenwald, I. S. & Horvitz, H. R. *Genetics* **96**, 147–164 (1980).
38. Capon, D. J. *et al. Nature* **304**, 507–513 (1983).
39. McCabe, P. C. in *PCR Protocols* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 76–83 (Academic, San Diego, 1990).

ACKNOWLEDGEMENTS. We thank M. Han and P. Sternberg for communicating their results before publication and for providing oligonucleotides for DNA sequencing of *let-60 ras* mutations. We also thank D. Clark and D. Baillie for *let-60* recessive lethal and deficiency strains, D. Parry and S. K. Kim for the *n1700* mutant, M. Han, G. Jongeward and P. Sternberg for the *sy103* and *sy130* mutants, S. C. Kim for the duplication *nDp5* and M. Stern for help with the laser beam microsurgery experiments. We thank C. Bargmann, J. Kaplan, L. DeLong, S. K. Kim and other members of the Horvitz and Meyer laboratories for discussions. This work was supported by the United States Public Health Service. G.J.B. is a Predoctoral Fellow of the Howard Hughes Medical Institute. H.R.H. is an Investigator of the Howard Hughes Medical Institute.

Primary structure of *Torpedo marmorata* chloride channel isolated by expression cloning in *Xenopus* oocytes

Thomas J. Jentsch, Klaus Steinmeyer & Gisela Schwarz

Centre for Molecular Neurobiology (ZMNH), Hamburg University, Martinistrasse 52, D-2000 Hamburg 20, Germany

A complementary DNA encoding a voltage-gated chloride channel from *Torpedo marmorata* electric organ was cloned by expressing hybrid-depleted messenger RNA in *Xenopus* oocytes. The predicted protein has a sequence of 805 amino acids containing several putative membrane-spanning domains. Expression of the protein in *Xenopus* oocytes shows that it is sufficient for channel function.

CHLORIDE channels are present in the plasma membranes of most cells and participate in various cellular functions, such as regulation of cell volume¹, transepithelial transport^{2,3} and stabilization of membrane potential in muscle⁴. There is a large variety of different Cl⁻ channels^{2,5}, and their importance is underscored by diseases such as cystic fibrosis, where mutations in the cystic fibrosis gene product⁶ block the cyclic AMP activation of epithelial Cl⁻ channels (for reviews, see refs 5, 7). Further, in certain forms of myotonia a decrease in plasma membrane Cl⁻ conductance leads to an increase in excitability⁸.

Identification and purification of Cl⁻ channels has been hampered by the absence of high-affinity ligands, with the exception of those binding to postsynaptic glycine receptors and GABA receptors. These ligand-gated Cl⁻ channels, which have been cloned^{9,10}, probably belong to a different class of channels. Recently, putative Cl⁻-channel proteins from kidney and trachea have been partially purified¹¹, but it is unclear which of the main proteins (of relative molecular mass 97K, 64K, 40K and 27K) represents a Cl⁻ channel. Monoclonal antibodies raised against *Necturus* gallbladder cells have been reported to inhibit Cl⁻ conductance¹². One of these antibodies reacts against proteins of *M_r* 219K and 69K, which have been suggested to be components of a Cl⁻ channel.

We used the electric ray *Torpedo* as a model system to isolate a Cl⁻ channel. The electric organ of *Torpedo californica* is an abundant source of a voltage-gated Cl⁻ channel (refs 13, 14; for review see ref. 15). This channel is thought to ensure the high conductance of the non-innervated membrane of the electrocyte necessary for efficient current generation caused by Na⁺ influx through the acetylcholine receptor at the innervated membrane. Previous attempts¹⁶ (including our own¹⁷) to identify the channel protein by using chemicals that block the *Torpedo* channel¹³ and presumably bind covalently to it, have failed because of the low binding specificity of the inhibitors used. In the absence of biochemical data on the protein(s) constituting this channel, we decided to rely only on function by using an expression cloning approach. By contrast to other Cl⁻ channels¹⁸, the channel from *Torpedo* electric organ can be expressed easily in *Xenopus* oocytes^{17,19}. We have now cloned the complete cDNA encoding this channel and show that a single protein subunit of calculated *M_r* ~89K, which contains several putative membrane-spanning domains, is sufficient for functional expression in *Xenopus* oocytes.

Size of active mRNA

A Cl⁻ channel from *T. californica* electric organ has been

expressed in *Xenopus* oocytes^{17,19}. We confirmed that a similar Cl⁻ conductance is expressed using RNA from *T. marmorata*, a closely related species. Consistent with the presumed abundance of the channel, small amounts (~5 ng) of total, non-poly(A)⁺ selected RNA were sufficient to cause a fivefold increase in oocyte plasma membrane conductance 2 days after injection. As described for *T. californica*¹⁹, the newly expressed channel is not calcium-dependent, unlike the endogenous *Xenopus* oocyte Cl⁻ channel(s)^{20,21}.

We determined the size of the mRNA(s) responsible for channel expression by size-fractionation using denaturing gel electrophoresis. A single mRNA size class of 9–10 kilobases (kb) is sufficient for functional expression in oocytes (Fig. 1). Mixing RNA from that fraction with others of lower sizes does not cause a further increase of membrane conductance. The difference in size to that published previously (~6 kb)¹⁹ for *T. californica* may result from species differences or from the use of non-denaturing sucrose gradients¹⁹.

Complementary DNA cloning

Our results are compatible with a single protein subunit being sufficient for channel function. A 9-kb mRNA may (but need not) encode a very large protein. In that case, a direct expression cloning approach using RNA transcribed *in vitro* from cDNA clones^{22,23} is difficult because of technical limitations in obtaining complete cDNA. We therefore chose a hybrid-depletion approach²⁴ in which cDNA size is not critical and which can also succeed if the channel is a complex of several essential subunits.

A tightly size-selected, functionally active mRNA fraction (Fig. 1) was used to construct a cDNA library in a plasmid vector. Single-stranded DNA derived²⁵ from small groups of clones was used to deplete the corresponding mRNAs from total *Torpedo* electric organ RNA which was then injected into *Xenopus* oocytes. A group of clones containing the desired cDNA should then cause attenuation or disappearance of Cl⁻-channel expression. The current flowing through nicotinic acetylcholine receptors, which are efficiently co-expressed by electric organ RNA¹⁹, was used as an internal control (Fig. 2b–e). After screening about 3,000 clones, we identified two groups of clones which reproducibly reduce Cl⁻-channel expression without interfering with acetylcholine receptor function (see Fig. 2c). Both groups contain 2–3 clones attenuating Cl⁻ currents, several of which have repetitive DNA sequences, recognize a broad smear on northern blots and probably exert nonspecific effects (by hybridizing to repetitive RNA sequences in untranslated regions of a large subset of the RNA population). Another clone (DC10) suppressing Cl⁻-channel expression (Fig. 2e), however, has a cDNA insert of 305 base pairs (bp) with a single open reading frame and no repetitive sequences. This insert hybridizes to a band of 9–10 kb in northern blots, which is most abundant in electric organ. We used this clone as a probe to isolate larger clones from a second cDNA library. From the number of clones detected, we estimate the corresponding mRNA to be abundant (≥0.1%), consistent with the electric organ being an abundant source of Cl⁻ channels. This probably implies that we have repeatedly overlooked correct clones during

the screening procedure. Indeed, the degree of functional depletion achieved with the initial clone is rather weak (Fig. 2e). Non-overlapping parts from clones identified by hybridization with clone DC10 were then used in additional hybrid-depletion assays and hybrid-arrest experiments using anti-sense oligonucleotides²⁶. This always led to a specific inhibition (by 70 to 90%) of Cl^- -channel expression, excluding non-specific hybridization.

Our hybrid depletion approach could result in the cloning of a protein other than the channel itself—an enzyme involved in some essential post-translational processing, for example. Thus, we had to show that the isolated clone indeed encodes the channel, or at least a subunit of a possible channel complex.

Channel expression

Three clones with cDNA inserts ranging from 2.7 to 3.2 kb were

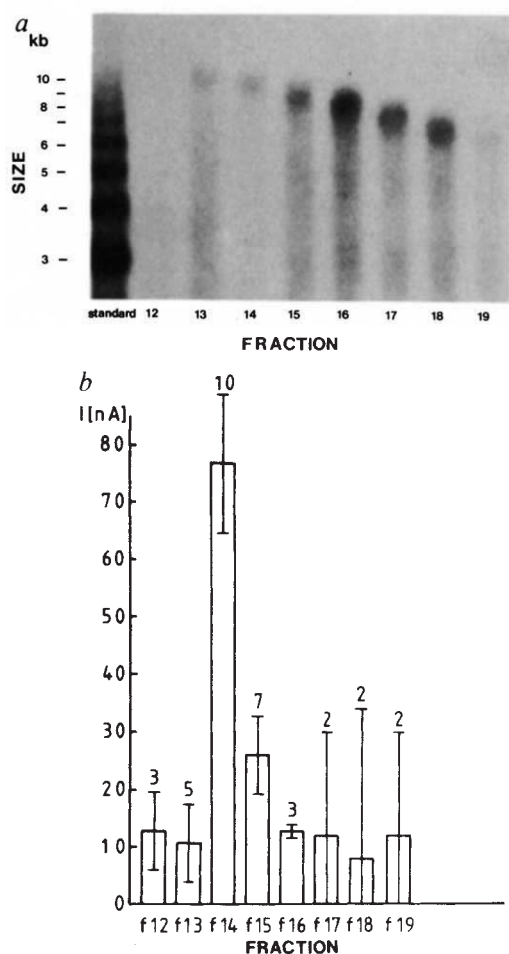


FIG. 1 Functional expression of size-fractionated mRNA. *a*, Northern blot analysis of size-fractionated RNA probed with labelled oligo(dT) to detect all mRNAs; *b*, Functional expression of the mRNA fractions shown in *a* in *Xenopus* oocytes. The mean differential current necessary to clamp the oocytes from -80 to -10 mV, from which the corresponding current of non-injected control oocytes has been subtracted, is plotted against the fraction number. Numbers above the columns, number of oocytes measured; bars, s.e.m.

METHODS. About $60 \mu\text{g}$ of poly(A)⁺-selected RNA from *T. marmorata* electric organ was size-fractionated by denaturing electrophoresis on a 0.8% agarose gel as described²⁴. The size of the resulting fractions was analysed by northern blot analysis using ³²P-labelled oligo(dT) as a probe²⁴. The size standard is a denatured end-labelled 1-kb DNA ladder (BRL). About 1/600 of the RNA from single fractions was injected into *X. laevis* oocytes which were obtained and manipulated as described⁴². Current-voltage relationships of injected and control oocytes were obtained by two-electrode voltage clamp 2 days after injection. The command voltage was varied between -100 and 0 mV using the program shown in Fig. 2 (inset).

isolated from a random-primed cDNA library using the initial clones as probes. Partial sequence analysis (showing a candidate initiator ATG at the 5' end and stop codons in all reading frames at the 3' end) indicated that they contain the complete coding region, so we transcribed RNA *in vitro* from all three of them. Injection of any of these RNAs into oocytes causes the expression of a large Cl^- conductance (Fig. 2f) similar to that observed with injection of total *Torpedo* electric organ RNA (Fig. 2b), demonstrating that a single polypeptide encoded by this cDNA is sufficient for Cl^- -channel activity.

Ion-substitution experiments show that the increase in conductance is carried by chloride ions (Fig. 2f). Current-voltage relationships were measured (using the command voltage program of Fig. 2, inset), and the hyperpolarization induced activation of the channel resulted in an outwardly rectifying current similar to the one elicited by total RNA (Fig. 3). The resting membrane potential ($I = 0$) is shifted towards the Cl^- -equilibrium potential of *Xenopus* oocytes (-20 to -30 mV)²⁰. When extracellular Cl^- concentration is decreased, the resting potential shifts further towards the new, more positive Cl^- -equilibrium potential. Currents are reduced, especially at hyperpolarizing voltages, presumably because Cl^- -influx is limited by reduced availability of chloride (half-saturation for the *T. californica* channel occurs at 75 mM Cl^-)¹⁵. Further, 1 mM DIDS (4,4'-diisothiocyanato-stilbene-2,2'-disulphonic acid) and 1 mM diphenylamine-2-carboxylate²⁷ (DPC), known inhibitors of Cl^- channels, cause a large inhibition of this conductance and shift the resting potential back to values close to those of uninjected controls. This result was observed for channels expressed either from total RNA or from *in vitro* RNA. The effect of DPC was fully reversible after wash-out, but as described¹³ for the reconstituted *T. californica* channel, the inhibition by DIDS is irreversible, presumably because of covalent binding. In contrast to the reconstituted channel, however, low concentrations of DIDS ($\leq 10^{-4} \text{ M}$) are ineffective within a few minutes. The chloride conductance induced by *in vitro* RNA is not dependent on either extra- or intracellular Ca^{2+} , as oocytes injected with EGTA and measured in Ca^{2+} -free saline containing 1 mM EGTA also display the typical chloride conductance (our unpublished results). This virtually excludes the possibility that we have cloned an activator of the endogenous oocyte Ca^{2+} -dependent Cl^- channel^{20,21}.

The time- and voltage-dependence of currents induced by cRNA was determined by stepping the voltage for 25 s to values between 0 and -100 mV from either a depolarizing holding potential (-10 mV; Fig. 3c), or from a hyperpolarizing potential (-100 mV; Fig. 3d). Starting from -10 mV, a voltage-induced activation of currents is observed for voltages more negative than about -40 mV (Fig. 3c). This activation is slow (of the order of several seconds), and the tail-currents indicate that subsequent inactivation after stepping back to -10 mV is even slower. By stepping from a holding potential of -100 mV to more positive voltages, we determined the time-dependent current-voltage relationship of channels activated by hyperpolarization (Fig. 3d). Depolarizing to -90 or -80 mV induces only small currents, which then increase with larger depolarizations. This leads to an overall outward rectification of currents after hyperpolarization-induced activation (Fig. 3a, b). These latter current-voltage relationships were obtained after passing several times through the staircase program shown in Fig. 2 (inset), which activates the channel by its hyperpolarizing steps.

Chloride conductance expressed from total electric organ RNA shows the same complex voltage-dependence (our unpublished results). The channel is slowly activated by hyperpolarization. Once activated, current flow increases with depolarization, leading to an apparent outward rectification. This compares well with the characteristics of the reconstituted *T. californica* channel (see ref. 15). This channel is activated slowly (within seconds) when the *cis* side of the bilayer is made negative. Once activated, however, the probability of the channel being open rapidly

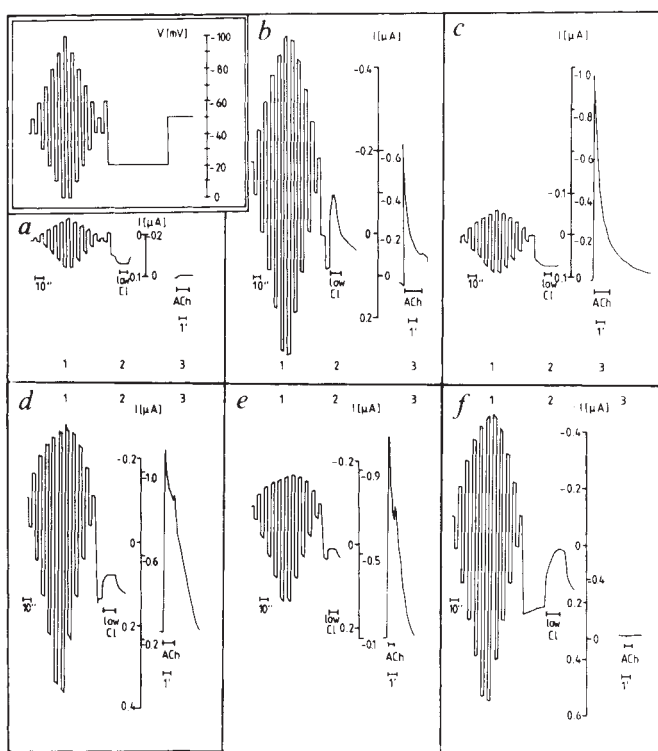


FIG. 2 Expression cloning assays. During the screening procedure, three tests were performed on each oocyte using two-electrode voltage-clamp technique: (1) determination of current-voltage curves by varying the command voltage in steps (4 s each) of 10 mV between -100 and 0 mV (inset, top left). (2) Direct determination of Cl^- current induced by changing bath $[\text{Cl}^-]$ from 104 to 7 mM (replaced by cycamate; bars) at $V = -20$ mV (Cl^- current is larger at depolarized V). (3) Measurement of current elicited by 1 mM acetylcholine (ACh) (in the presence of 10^{-5} M atropine to block endogenous muscarinic receptors; bars) at $V = -50$ mV. Tests are indicated by numbers in all panels. Note different current and timescales. Inset, Command voltage program used during voltage clamp experiments, chosen to yield an easily recognizable 'fingerprint' of channel activity during the screening procedure. **a**, Uninjected control oocyte with current-voltage relationship typical for native oocytes: lowering $[\text{Cl}^-]$ causes no detectable Cl^- current (the small current is caused by liquid junction potentials⁴³); acetylcholine has no effect. **b**, Oocyte injected with ~ 10 ng total electric organ RNA. An outward-rectifying current typical for the *Torpedo* Cl^- -channel is observed: reducing bath $[\text{Cl}^-]$ leads to an inward current (underestimated owing to the presence of liquid junction potentials⁴³); application of acetylcholine leads to a large, inactivating inward current. **c**, Oocyte injected with electric organ RNA depleted with $60 \mu\text{g}$ single-stranded DNA (ssDNA) mixture from group DC (12 clones), the positive group finally identified; expression of Cl^- conductance is largely reduced (tests 1 and 2), while the acetylcholine-induced current is fully present. **d**, Oocyte injected with RNA depleted with $\sim 5 \mu\text{g}$ single-stranded (ss) DNA derived from DC8, a single, negative clone from group DC: both Cl^- channels and acetylcholine receptors are fully expressed. **e**, Oocyte injected with RNA depleted with $\sim 5 \mu\text{g}$ ssDNA derived from DC10, the final positive clone from group DC: reduction of Cl^- current and full response to acetylcholine. We typically used $5 \mu\text{g}$ ssDNA when screening clones in groups of 12; larger inhibition (80–90%) of Cl^- currents was obtained when hybrid-depleting with $50 \mu\text{g}$ ssDNA from clone DC10. **f**, Oocyte injected with cRNA from 7134, a full-length clone identified by hybridization with clone DC10. Large Cl^- currents are observed, and no response to acetylcholine. **METHODS:** Random primed cDNA⁴⁴ was prepared from the active mRNA fraction 14 of Fig. 1, ligated to adaptors compatible with the *Bst*XI overhang in CDM8 (ref. 45), size selected, and ligated into *Bst*XI-cut plasmid vector pBst. This vector was constructed by deleting the *Bst*XI site in pBluescriptKS(+) (Stratagene) and inserting the stuffer fragment from CDM8 (ref. 45) into the single *Xho*I site. After transformation of *Escherichia coli* XL1-blue (Stratagene) about 1.2×10^6 clones were obtained. A second library in pBst was made by ligating size-selected cDNA (obtained by random-primed synthesis with RNaseH-free Moloney murine leukaemia virus reverse transcriptase (BRL) from $1.0 \mu\text{g}$ non-fractionated poly(A)⁺ RNA) into pBst. Transformation of *E. coli* XL1-blue by electroporation gave about 6×10^6 primary transformants. The first library was screened by hybrid-depletion assays and the second one with radiolabelled cDNA fragments. For hybrid depletion, ssDNA was prepared²⁵ from groups of 12 clones ($\sim 60 \mu\text{g}$ total) and used with $20 \mu\text{g}$ total electric organ RNA as described²⁴. The single-stranded DNA was not linearized and no poly(dA) was included. The final depleted RNA was taken up in $8 \mu\text{l}$ H_2O , 50 nl of which were injected into each oocyte. Capped *in vitro* RNA was synthesized using T3 RNA polymerase from 7134 after linearizing with *Xba*I. About 0.5 ng RNA were injected per oocyte, which were measured after 1–2 days in ND96 saline (ref. 42) at room temperature.

increases with more positive voltage on the *cis* side. Assuming that the cloned and the reconstituted channels are identical, this suggests that the *cis* side of the bilayer corresponds to the cytoplasm. This could explain the lack of an extracellular effect of micromolar concentrations of DIDS in our system: the reconstituted channel is inhibited from the *cis* side¹³.

Protein structure

Figure 4 shows the 2,674-nucleotide sequence of the cDNA insert of clone 7134, one of the clones eliciting chloride conduct-

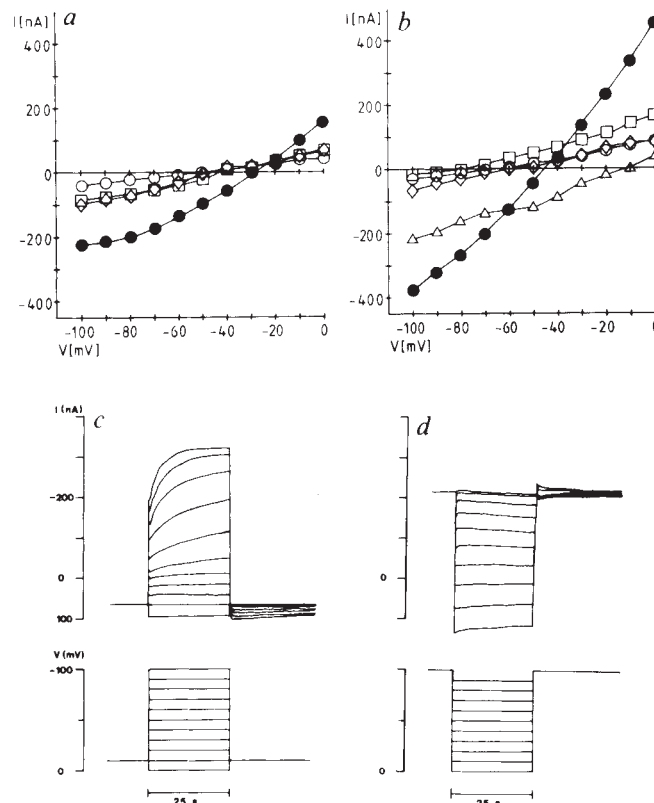


FIG. 3 Electrophysiological characterization of chloride conductance expressed from total RNA (**a**) or cRNA (**b**, **c**, **d**). **a**, **b**, Current-voltage relationships of chloride conductance expressed from total electric organ RNA (**a**) or cRNA derived from chloride channel clone (**b**). The current necessary to clamp the oocytes to the indicated voltages is shown for: (○) uninjected controls; (●) oocytes injected with total (**a**) or synthetic mRNA (**b**), measured in ND96 saline. (◇) RNA-injected oocytes, measured in ND96 containing 1 mM DIDS. (□) RNA-injected oocytes, measured in ND96 containing 1 mM diphenylamine carboxylate. (Δ) cRNA-injected oocyte, measured in low-chloride saline (96 mM NaCl replaced by sodium cycamate in ND96). **c**, **d**, Time- and voltage-dependence of chloride currents expressed from cRNA. In **c**, oocyte potential was held depolarized at -10 mV for ≥ 4 min and then stepped for 25 s to the values indicated at the bottom. This led to a slow activation of currents for potentials more negative than -40 mV. In **d**, the conductance was kept activated by holding for ≥ 1 min at -100 mV, and the potential was then stepped for 25 s to more positive values. This demonstrates the apparent outward rectification of macroscopic currents of the channel activated by hyperpolarization.

METHODS. Current-voltage relationships (**a**, **b**) were determined using two-electrode voltage clamp two days after injecting ~ 10 ng total electric organ mRNA (**a**) or ~ 0.5 ng synthetic mRNA derived from clone 7134 (**b**). The command voltage was varied according to the program in Fig. 2. Results obtained from single, representative oocytes are shown. As DIDS acts irreversibly, it was applied last. The resting membrane potentials of the batch of oocytes used for **b** were more negative than those of **a**, explaining the shift of the curves towards more negative values of V . This may also explain why the reversal potential observed with cRNA injected oocytes (**b**) is more negative than the one in **a**, as in the presence of a larger K^+ conductance expressed in parallel with a large Cl^- conductance intracellular Cl^- concentration will equilibrate to yield a more negative equilibrium potential. When liquid junction potentials⁴³ are taken into account, the curve for low-chloride medium (Δ) will be shifted by ~ 10 mV towards more positive values. The fact that in **b** the resting potential in the presence of DPC is slightly hyperpolarized versus controls might be due to inhibition of unspecific cation channels by DPC (ref. 2). In **c** and **d**, conventional two-electrode voltage clamp was performed on oocytes (from a different batch) one day after injection of capped *in vitro* RNA (~ 0.4 ng) from clone 7134.

513

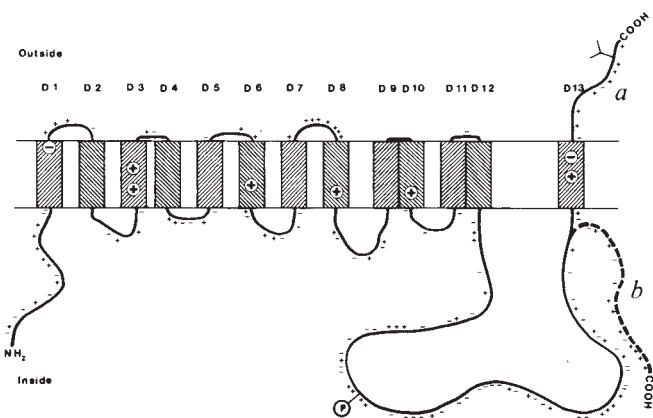


FIG. 6 Proposed membrane topology for the *Torpedo* chloride channel. For *a*, it is assumed that all hydrophobic domains indicated in Fig. 5 cross the plasma membrane, and that the N-terminal part of the protein is located intracellularly. The potential glycosylation site at amino-acid 801 could then be used (Y), as could be the phosphorylation site at 600 (P). *b*, domain D13, which (with D3 and D8) displays intermediate hydrophobicity, is assumed not to cross the plasma membrane, leading to a model with 12 membrane spans. Strongly charged amino acids (Arg, Lys, Glu, Asp) are indicated by their charges.

hydrophilic pore. One such alternative is shown in Fig. 6*b*, in which D13 is assumed not to cross the membrane. In this case, 12 membrane spans would be predicted, with both N and C termini on the cytoplasmic side.

Several putative membrane-spanning domains contain residues with charged side chains. Although D1 and D13 contain negative charges, positive charges (Arg or Lys) are present in domains D3, D6, D8, D10 and D13, and histidines are found in D4, D9, D11 and D12. This could provide a selectivity filter for chloride ions, possibly in conjunction with positive charges just next to putative membrane-spanning domains such as the highly positively charged region between D7 and D8.

Discussion

Expression from cRNA proves that a single polypeptide is sufficient for the function of the *Torpedo* chloride channel, while our hybrid-depletion and hybrid-arrest experiments suggest that we have cloned the major chloride conductive pathway from the electric organ which can be expressed in *Xenopus* oocytes. The protein has up to 13 putative membrane-spanning domains, which are considered to be a hallmark of ion channels and other transport proteins. Database searches show that it is a novel protein while matrix comparisons³² revealed no significant simi-

ilarity to glycine¹⁰ and GABA receptors⁹, which are ligand-gated Cl⁻ channels or band 3 (ref. 48), the electroneutral anion exchanger of erythrocytes.

Many cloned cation-selective channels probably belong to a superfamily of ion channels. Members of a large gene family of potassium channels originally identified³³⁻³⁵ as transcripts of the *Drosophila shaker* locus have six putative membrane-spanning domains, whereas the voltage-dependent Na-channel³⁶ and the dihydropyridine-sensitive Ca-channel³⁷ have four repeats of six such domains. The cGMP-gated cation channel³⁸ of the retinal rod photoreceptor may also belong to the same superfamily³⁹. The chloride channel described here does not fit easily into this group. Although omission of one of the more questionable membrane-spanning domains (such as D13; Fig. 6*b*) may give two times six membrane spans, there are no clear equivalents to the charged S4 domains (repeats of Arg-X-X, where X are mainly hydrophobic amino acids, and Arg may be replaced by Lys) present in this superfamily. This segment is thought to provide the voltage-sensor. In view of the presence of a similar segment in the cGMP-gated channel, which is only weakly voltage-sensitive, other interpretations are possible³⁹, and an entirely different voltage-sensitive potassium channel⁴⁰ has no such segment. Although the channel described here is voltage-dependent, only one stretch (positions 75-81) is at all similar to a S4 segment. This stretch, in being located after the first transmembrane segment, is atypical. Voltage-dependence could be caused by charges found in several putative transmembrane domains of the *Torpedo* channel.

We cannot state unequivocally whether the Cl⁻ channel we have cloned is the same as the one studied in the lipid-bilayer system¹³⁻¹⁵. The chloride selectivity of the present channel, its voltage-dependence (slow activation by hyperpolarization and, once activated, increase in conductivity with depolarization), together with the abundance of its mRNA, seem to argue for such an identity. The situation will be clarified by single-channel analysis of channels expressed from the present cDNA.

On the basis of biophysical data, the *Torpedo* chloride channel as observed in reconstitution was suggested¹⁴ to be a homodimer (named 'double-barreled channel'). This suggestion could be clarified by crosslinking studies. A similar 'double-barreled' Cl⁻ channel has recently been described⁴¹ in the mammalian kidney which displays nearly the same peculiar voltage-activation as that described here on the macroscopic level. These findings, together with the presence of a potential cAMP-dependent phosphorylation site in the Cl⁻ channel reported here, may increase our chances of isolating the epithelial, cAMP-regulated Cl⁻ channel involved in cystic fibrosis using our cDNA as a probe. □

Received 5 September; accepted 25 October 1990.

- Worrell, R. T., Butt, A. G., Cliff, W. H. & Frizzell, R. A. *Am. J. Physiol.* **256**, C1111-C1119 (1989).
- Gögelein, H. *Biochim. biophys. Acta* **947**, 521-547 (1988).
- Liedtke, C. M. *A. Rev. Physiol.* **51**, 143-160 (1989).
- Bretag, A. H. *Physiol. Rev.* **67**, 618-724 (1987).
- Frizzell, R. A. *Trends Neurosci.* **10**, 190-193 (1987).
- Riordan, J. R. *et al. Science* **245**, 1066-1073 (1989).
- Welsh, M. J. *FASEB J.* **4**, 2718-2725 (1990).
- Rüdel, R. & Lehmann-Horn, F. *Physiol. Rev.* **65**, 310-356 (1985).
- Schofield, P. R. *et al. Nature* **328**, 221-227 (1987).
- Grenningloh, G. *et al. Nature* **328**, 215-220 (1987).
- Landry, D. W. *et al. Science* **244**, 1469-1472 (1989).
- Finn, A. L., Tsai, L.-M., Falk, R. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7649-7652 (1989).
- White, M. M. & Miller, C. *J. biol. Chem.* **254**, 10161-10166 (1979).
- Miller, C. & White, M. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2772-2775 (1984).
- Miller, C. & Richard, E. A. in *Chloride Channels and Carriers in Nerve, Muscle and Glial Cells* (eds Alvarez-Leefmans, F. J., & Russel, J. M.) 383-405 (Plenum, New York 1990).
- Taguchi, T. & Kasai, M. *Biochem. biophys. Res. Commun.* **96**, 1088-1094 (1980).
- Jentsch, T. J., Garcia, A. M. & Lodish, H. F. *Biochem. J.* **261**, 155-166 (1989).
- Kroll, B. *et al. Am. J. Physiol.* **257**, L284-L288 (1989).
- Sumikawa, K., Parker, I. & Miledi, R. *EMBO J.* **3**, 2291-2294 (1984).
- Barish, M. E. *J. Physiol., Lond.* **342**, 309-325 (1983).
- Boton, R., Singer, D., & Dascal, N. *Pflügers Arch. Eur. J. Physiol.* **416**, 1-6 (1990).
- Hediger, M., Coady, M. J., Ikeda, T. S. & Wright, E. M. *Nature* **330**, 379-381 (1987).
- Masu, Y. *et al. Nature* **329**, 836-838 (1987).
- Lübbert, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4332-4336 (1987).
- Vieira, J. & Messing, J. *Meth. Enzym.* **153**, 3-11 (1987).
- Kawasaki, E. S. *Nucleic Acids Res.* **13**, 4991-5004 (1985).

- Distefano, A. *et al. Pflügers Arch. Eur. J. Physiol.* **405**, S95-S100 (1985).
- Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
- Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
- Chou, P. Y. & Fasman, G. D. *A. Rev. Biochem.* **47**, 251-276 (1978).
- Edelman, A. M., Blumenthal, D. K. & Krebs, E. G. *A. Rev. Biochem.* **56**, 567-613 (1987).
- Staden, R. *Nucleic Acids Res.* **10**, 2951-2961 (1982).
- Papazian, D. M., Schwarz, T. L., Tempel, B. L., Jan, Y. N. & Jan, L. Y. *Science* **237**, 749-753 (1987).
- Baumann, A. *et al. EMBO J.* **6**, 3419-3429 (1987).
- Kamb, A., Iverson, L. E. J. & Tanouye, M. A. *Cell* **50**, 405-413 (1987).
- Noda, M. *et al. Nature* **312**, 121-127 (1984).
- Tanabe, T. *et al. Nature* **328**, 313-318 (1987).
- Kaup, U. B. *et al. Nature* **342**, 762-766 (1989).
- Jan, L. Y. & Jan, Y. N. *Nature* **345**, 672 (1990).
- Takumi, T., Ohkubo, H. & Nakanishi, S. *Science* **242**, 1042-1045 (1988).
- Sansom, S. C., La, B.-Q. & Carosi, S. L. *Am. J. Physiol.* **259**, F46-F52 (1990).
- Colman, A. in *Translation and Transcription* (eds Hames, B. D. & Higgins, S. J.) 271-302 (IRL, Oxford, 1984).
- Jentsch, T. J., Matthes, H., Keller, S. K. & Wiederholt, M. *Pflügers Arch. Eur. J. Physiol.* **403**, 175-185 (1985).
- Gubler, U. & Hofman, B. J. *Gene* **25**, 263-269 (1983).
- Seed, B. *Nature* **329**, 840-842 (1987).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
- Kopito, R. R. & Lodish, H. F. *Nature* **316**, 234-238 (1985).

ACKNOWLEDGEMENTS. We thank A. Block and W. Hamel for help with preparation of single-stranded DNA. Supported by the Bundesministerium für Forschung und Technologie (BMFT) and the Cystic Fibrosis Foundation.

Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A

Kiyoshi Nagai, Chris Oubridge, Timm H. Jessen, Jade Li & Philip R. Evans

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The crystal structure of the RNA binding domain of the U1 small nuclear ribonucleoprotein A, which forms part of the ribonucleoprotein complex involved in the excision of introns, has been solved. It contains a four-stranded β sheet and two α helices. The highly conserved segments designated RNP1 and RNP2 lie side by side on the middle two β strands. U1 RNA binding studies of mutant proteins suggest that the RNA binds to the four-stranded β sheet and to the flexible loops on one end.

MANY RNA-binding proteins are involved in the splicing, storage and transport of RNA, and in this way play crucial parts in regulating gene expression. In the nucleoplasm newly transcribed RNA molecules are capped and polyadenylated, and form ribonucleoprotein particles when complexed with proteins that bind heterogeneous nuclear RNA (hnRNA) or poly(A)¹. These pre-messenger-RNA molecules then interact with a series of small nuclear ribonucleoprotein particles (snRNPs) so that introns are spliced out to give mature messenger RNAs²⁻⁴. Mature mRNAs are exported through the nuclear pores to the cytoplasm as ribonucleoprotein particles.

Many RNA-binding proteins involved in these processes contain single or multiple copies of a homologous 90-amino-acid sequence. This RNP (ribonucleoprotein) motif is thought to form an RNA-binding domain^{5,6}. The overall sequence similarity of this motif in over 20 different proteins is low; nevertheless their evolutionary relationship is evident from the presence of two highly conserved short segments in this motif, referred to as RNP1 (or RNP octamer consensus)^{7,8} and RNP2 (ref. 5) (Fig. 1). The RNP motif is considered to be an RNA-binding module which has evolved to bind different RNA sequences and which, through gene duplication and with auxiliary domains, has given rise to RNA-binding proteins with diverse functions^{5,9}.

The first steps of pre-mRNA splicing involve the binding of U1 and U2 snRNPs to the 5' exon-intron junction and branch site of pre-mRNA, respectively²⁻⁴. Two proteins in human U1 snRNP, designated U1 70K and U1 A proteins, contain the RNP motif^{10,11}. In these two cases a polypeptide slightly larger than the RNP motif does have specific RNA-binding activities^{6,12-14}. The U1 70K and A proteins bind to stem-loops I and II of U1 snRNA respectively^{6,12-16}. U2 B' protein is a component of U2 snRNP and binds to stem-loop IV of U2 snRNA only in the presence of U2 A' protein^{13,17}. U1 A and U2 B' proteins are closely related and both have two copies of the RNP motif connected by polypeptides of different lengths^{11,18}. The N-terminal RNP domains of U1 A and U2 B' proteins show 77% sequence identity, and their C-terminal domains show 88% identity. Only the N-terminal domains of U1 A and U2 B' proteins are necessary for specific binding to their cognate RNA molecules^{12,13,17}. In this region these two proteins differ only at 20 out of 102 positions and some of these differences are involved in differential RNA and protein binding^{13,17}.

We have solved the crystal structure of the N-terminal RNP domain of the U1 A protein at 2.8 Å resolution. We have made mutants of this domain, and studied the U1 RNA-binding ability of the mutant proteins. This has allowed us to identify some of the residues involved in RNA-binding. Here we describe the crystal structure of the domain and also suggest how it may interact with RNA.

Structural determination

Various segments of human U1 snRNP protein A were expressed in *Escherichia coli* using the T7 RNA polymerase expression vector¹⁹ and purified on a CM-Sepharose column. Crystals of the A102 protein, containing the N-terminal RNP domain residues 1-102, were grown under various conditions, but the best of these crystals diffracted only to 4.5 Å resolution. Our preliminary two-dimensional (2-D) NMR study of the A102 protein indicates that the C terminus is extremely mobile (D. Neuhaus, T.-H.J. and K.N., unpublished results). Better crystals were obtained with the shorter protein A95 (residues 1-95), grown by microdialysis against ~2 M potassium/sodium phosphate, pH 5.5. These crystals diffract to at least 2.2 Å resolution, and belong to the cubic space group I4₁32 with a cell edge of 148.1 Å, containing two molecules of the RNP domain in the asymmetric unit. Details of the structural determination are given in Table 1. Because the wild-type protein does not contain any cysteines to react with heavy metals, 10 mutant proteins were engineered with single polar side chains changed to cysteines. Heavy-atom derivatives were prepared by reaction with methylmercury nitrate in solution, and crystallization was attempted for both the reacted and unreacted proteins. Methylmercury derivatives of four of the mutants produced crystals isomorphous with the native, though for the mutant in which Ser 71 is replaced by Cys (S71C mutant; single-letter amino-acid code), better derivative crystals were obtained by soaking unreacted crystals in methylmercury (Table 1). The other six mutants produced either no crystals, different crystals or small crystals. Complete diffraction data sets were collected for the native protein and four mercury derivatives: crystals of the unreacted Q85C mutant (Q, Gln) were used as the native protein because they were larger than those from the wild-type protein. The structure was solved by conventional isomorphous replacement methods, combined with solvent flattening (Table 1). The map was easily interpretable, with no uncertainties in tracing the polypeptide chain, except for disorder in the chain termini and in a surface loop (residues 50-52) in one of the subunits (see below). The model has been refined sufficiently to confirm the essential features of the structure described here, but refinement will be continued when higher resolution data are available. We have also calculated a difference Fourier map between the true wild-type and the Q85C mutant protein at 3.3 Å resolution, which showed that the mutation causes no significant change of structure.

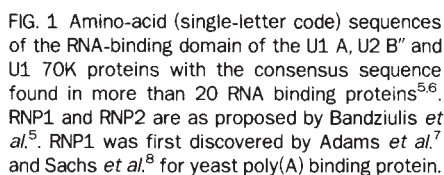
Structure of the RNP domain

Figure 2a is a representation of the secondary structure of the RNP domain of the U1 A protein. This domain consists of two α helices and a four-stranded antiparallel β sheet arranged in the order β 1-A helix- β 2- β 3-B helix- β 4 in the primary structure, and the four β strands are arranged in space in the order β 4- β 1- β 3- β 2 with two helices on the same side of the sheet.

Derivative	Resolution (Å)	Number of unique reflections	Mean redundancy	$R_{\text{merge}}^{\ddagger}$	Number of sites	$\langle F_{\text{H}} \rangle / E^{\S}$
Native (Q85C)	2.8	7,067	7.7	0.090	—	—
S71C + MeHg*	3.5	3,509	5.0	0.101	4	1.7
Q85C + MeHg†	2.8	6,913	8.0	0.082	4	1.2
Q38C + MeHg†	2.8	6,774	7.2	0.077	2	1.4
E61C + MeHg†	3.0	5,774	7.7	0.087	2	1.3

* Derivative prepared by soaking crystals of the mutant in 0.25 mM methyl mercury (MeHg) overnight.
† Derivatives crystallized from mutant proteins reacted with 1.5-fold excess MeHg for 1 h, then dialysed overnight against 50 mM NaCl.
‡ $R_{\text{merge}} = \sum |I_i - \langle I \rangle| / \sum \langle I \rangle$, where I_i is an individual measurement of intensity and $\langle I \rangle$ is the mean intensity for this reflection.
§ 'Phasing power', (r.m.s. calculated F_o)/(r.m.s. error), for centric reflections only.

Crystallographic studies of protein-DNA complexes have shown that Lys and Arg residues are crucial for binding to the phosphate backbone or bases of DNA³⁰. The A102 protein binds to U1 RNA as strongly as the whole U1 A protein (T.-H.J., C.O.



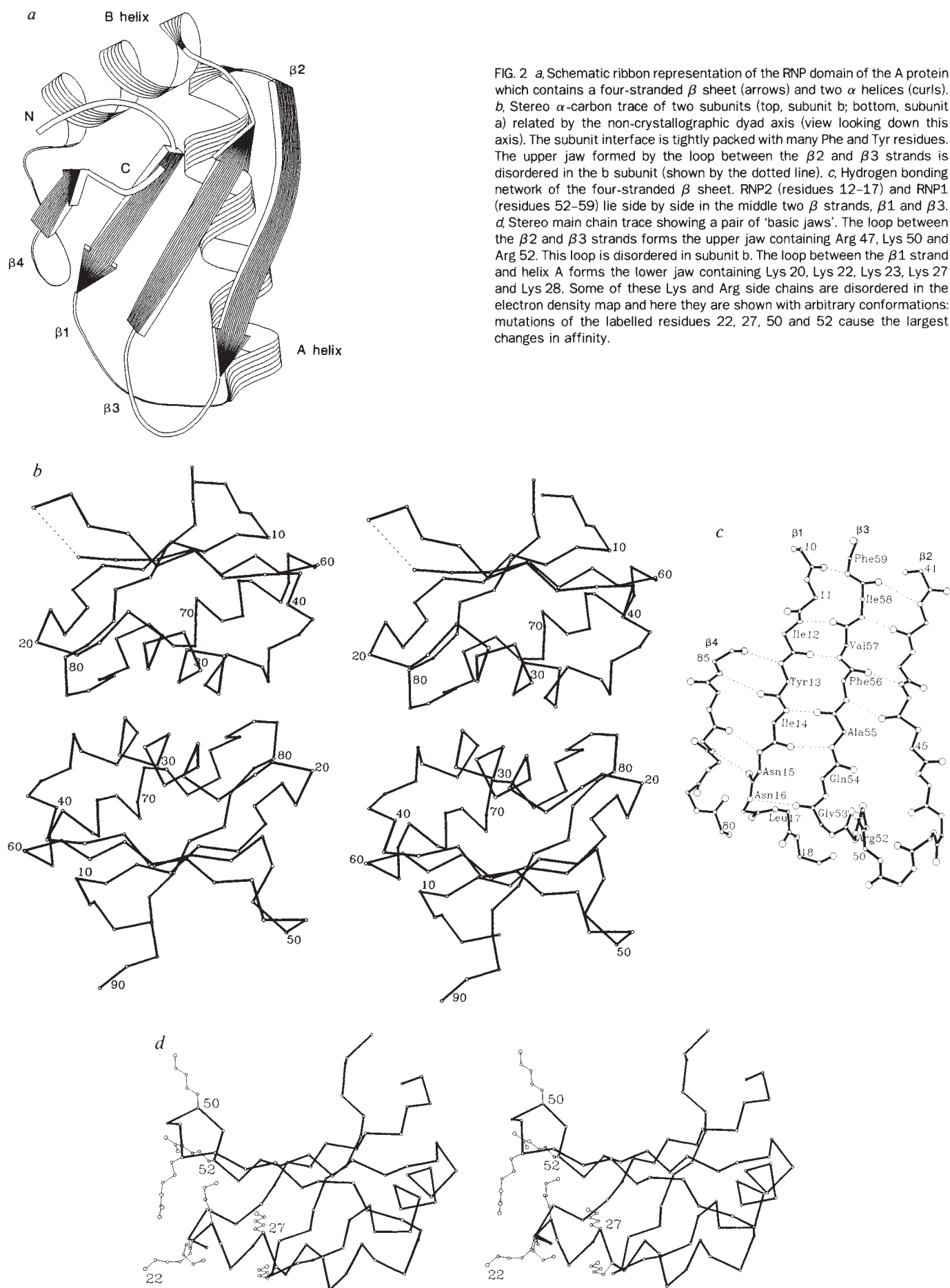


TABLE 2 Relative binding affinities of mutant A102 proteins to U1 RNA stem-loop II

Protein	Relative binding affinity	Residue in U2 B' protein
A102 wild type	+++	
R7Q	+++	R
K20Q	+++	K
K22Q	+	K
K23Q	++	K
K27Q	+	K
K28Q	+++	R
R47Q	++	K
K50Q	+	K
R52Q	—	R
K60Q	+++	K
R70Q	+++	R
K80Q	+	K
R83Q	++	R
K88Q	++	K
K96Q	+++	K
K98Q	+	R

The relative U1 RNA binding affinities of the mutant proteins are shown in comparison with the wild-type A102 protein (amino acids 1–102 of the U1 A protein). Total loss of specific binding is indicated by —; up to 10-fold and 100-fold reductions in binding strengths are indicated by ++ and +, respectively. U1 RNA stem-loop II (nucleotides 50–92) was synthesized by transcription *in vitro* with T3 RNA polymerase from linearized plasmids carrying the corresponding cDNA³¹. The RNA was labelled at the 5' end according to standard procedures and purified on denaturing polyacrylamide gels. The labelled RNA (~5 nM) was incubated with various amounts of the mutant proteins in 10 mM sodium-HEPES, pH 7.4, 50 mM KCl and 1 mM MgCl₂ at room temperature for 20 min. The complex was separated from unbound RNA on native 12% polyacrylamide gels containing 100 mM Tris borate, EDTA (pH 8.3) and 0.1% Triton X-100 (D. Bogenhagen *et al.*, manuscript in preparation); the wild-type protein was always used as the internal control. The gel was autoradiographed without any further treatment for 16 h at –70 °C, and the binding affinity estimated from the proportion of complexed and free RNA.

and K.N., unpublished results). It contains 11 Lys and 5 Arg residues, some of which are likely to interact with the backbone phosphates or bases of the RNA. We have replaced each Lys and Arg residue in turn by Gln, because it has an aliphatic chain with an uncharged hydrophilic head group. All the Lys and Arg residues are external and most of them are disordered in the electron density map, so these substitutions are unlikely to cause large structural perturbations. None of these substitutions caused any significant destabilization of the protein and all the mutants were indeed expressed in the folded form and could be purified to homogeneity as could the wild-type protein. We made the 46-nucleotide stem-loop II of U1 RNA by transcription *in vitro*, using phage T3 RNA polymerase³¹; we estimated the binding constants of the mutant protein by their shift in mobility on acrylamide gels. Table 2 shows the relative affinities of these proteins to the stem-loop II of U1 RNA. The wild-type protein binds to this RNA molecule with the apparent dissociation constant (K_d) of 1×10^{-8} M, which is significantly smaller than the previous estimate (8×10^{-8} M) using a rabbit reticulocyte *in vitro* translation system¹⁴. The specific RNA binding was abolished in the R52Q (mutation of Arg 52 to Gln) mutant. Reductions of 10–50-fold in the binding were observed with the K22Q, K27Q, K50Q, K80Q and K98Q mutants. The K23Q, R47Q, R83Q and K88Q mutants showed two- to fivefold reductions in the affinity; The R7Q, K20Q, K28Q, K60Q, R70Q and K96Q mutations had no significant effect on the U1 RNA binding affinity. Lutz-Freyermuth *et al.*¹⁴ reported that the A96 protein containing residues 1–96 has a U1 RNA-binding affinity comparable to that of the full-length A protein, but our experiments with purified proteins show that the K98Q mutant and the A96 protein have significantly reduced U1 RNA binding

affinity compared with the A102 protein (T.-H.J., C.O. and K.N., unpublished results).

RNA binding surface

Figure 4 shows a stereo view of the main-chain with all the Lys and Arg side chains coloured according to the size of the effect of the Gln substitutions on binding to U1 RNA stem-loop II. The effects are colour-coded in increasing order as pale blue, yellow, pink and red. Positively charged residues are also found at the equivalent positions in the U2 B' protein but it has the conservative substitutions K28R, R47K and K98R (Table 2)^{11,18}.

More than half of the Lys and Arg residues are located on the loops at one end of the four-stranded β sheet and most mutations (K20Q, K22Q, K23Q, K27Q, R47Q, K50Q, R52Q, K80Q and R83Q) in this region had a significant effect on the U1 RNA binding. The K88Q mutation, which also altered the U1 RNA binding, lies on the top left hand corner of the molecule as shown in Fig. 4. In the electron density map the C terminus of the molecule beyond residue 90 is marked by only weak density, suggesting it is disordered. A preliminary 2-D NMR study showed that the C terminus of the molecule is also highly mobile in the A102 protein. These two results indicate that the segment beyond residue 90 is not an integral part of the RNP domain, but part of the linker between the N- and C-terminal RNP domains. In fact, the C-terminal RNP domain of the U1 A and U2 B' proteins terminate at the residue equivalent to Thr 89 of the U1 A N-terminal domain (Fig. 1)^{11,18}. The linker peptide is extremely susceptible to proteases, suggesting that it is very flexible (T.-H.J., C.O. and K.N., unpublished observations). But a substantial reduction in RNA binding was observed on mutation or deletion of Lys 98, showing that this residue is essential for RNA binding and that this region of the molecule can become ordered on RNA binding, a common phenomenon for nucleic acid-binding proteins³⁰.

RNA binding specificities of U1 A and U2 B'

The U1 A protein binds to the stem-loop II of U1 RNA on its own, whereas the U2 B' protein binds to the stem-loop of U2 RNA only in the presence of U2 A' protein^{13,17}. The loop sequences of the U1 stem-loop II and the U2 stem-loop IV are AUUGCACUCC and (U)AUUGCAGUAC(C)(U) respectively. The binding specificities of these two stem-loops to the U1 A and U2 B' proteins can be interchanged by replacing the two underlined bases, showing that the exact loop size or stem sequence is not critical for binding¹³. This suggests that these two bases interact with amino-acid residues that differ between U1 A and U2 B' proteins. The amino-acid sequence of U1 A differs from that of U2 B' protein at only 20 out of the 102 positions^{11,18}. The segment between residues 40 and 49 of U1 A protein and the corresponding part of U2 B' protein are major determinants of their specificity for U1 and U2 RNA¹³. If we exclude internal residues and conserved residues in this region, L44V, S46L, R47K and S48T remain as candidates for the crucial differences in U2 B' that prevent binding of U1 RNA. The R47Q replacement has only a small effect on the binding of U1 RNA so this residue is unlikely to discriminate between U1 and U2 RNA; therefore S46, L44 and/or S48 may have important roles in distinguishing the U1 RNA stem-loop II and the U2 RNA stem-loop IV. It is possible that Ser 46 and/or Ser 48 form a hydrogen bond to cytosine at the seventh or ninth positions of the loop II of U1 RNA. In Fig. 4 these two Ser residues are coloured green, and external residues of RNP1 and RNP2, blue. All the residues critical to the U1 RNA binding are clustered about RNP1 and RNP2, nearly all on one face of the molecule.

The U2 A' protein is indispensable for the binding of the U2 B' protein to the U2 RNA-loop IV, and U2 A' forms a complex with U2 B' even in the absence of RNA^{13,17}. When residues 40–49 of U1 A were replaced by residues 37–46 of U2 B', that is, the β 2 strand and part of the β 2– β 3 loop, U2 A' binds weakly to U1 A; two additional mutations, D24E and K28R (on helix

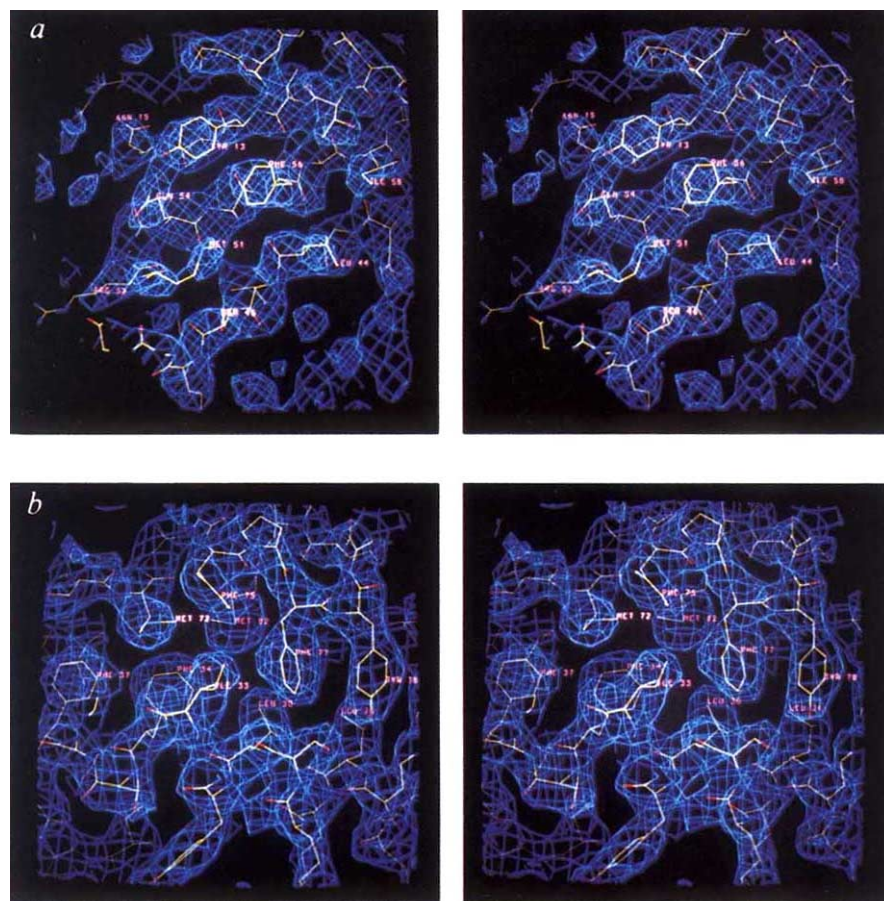


FIG. 3 The 2.8 Å resolution electron density map calculated with solvent-flattened, multiple isomorphous replacement (MIR) phases, at the contour level of $0.3 \text{ e}/\text{\AA}^3$. The atomic model fitted to the MIR map is also included²⁹. *a*, Region of the RNP1 and RNP2 segments. The $\beta 1$ strand containing RNP2, $\beta 3$ strand containing RNP1 and $\beta 2$ are clearly seen. *b*, RNP domain seen from the interface between two subunits related by the non-crystallographic dyad. This subunit interface is tightly packed with aromatic and large aliphatic residues.

A), are needed for strong binding¹⁷. It is likely from the structure that the U2 A' protein binds on the edge of the four-stranded β sheet by interacting with the $\beta 2$ strand and helix A (Fig. 4). Two bases in the U2 RNA loop IV differing from the corresponding positions in the U1 RNA loop II may interact with residues in U2 A' protein when it is bound to the edge of the $\beta 2$ strain.

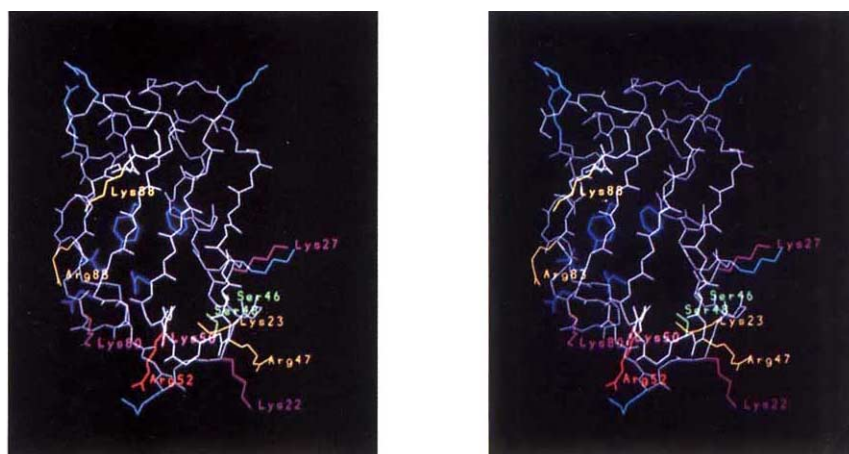
Discussion

The position of essential binding residues in the crystal structure shows that RNA interacts with the four-stranded β sheet containing the RNP1 and RNP2 consensus sequences and with the loops on one end of the sheet. What is the nature of this interaction between U1 RNA and U1 A protein? The loop II of U1 RNA contains the 10 unpaired nucleotides AUUGC-ACUCC, and this loop together with part of the stem will form

a structure which is comparable in size with the RNP domain. Models of RNA loops come from known structures of transfer RNA anticodon loops³²⁻³⁵. The crystal structure of glutamyl-tRNA synthetase/tRNA^{Gln} complex shows how the structure of this loop changes on binding synthetase³⁶. The anticodon loop of tRNAs contains seven unpaired nucleotides, and the sugar-phosphate backbone assumes a right angle between the second and third nucleotides of the loop. Bases before and after this turn are stacked on top of each other. In the structure of the complex, the three anticodon bases are splayed out and buried in the protein where they form extensive hydrogen bonds³⁶. This shows that the anticodon loop is flexible and undergoes large structural changes on interaction with the protein.

The loop II of U1 RNA and the loop IV of U2 RNA contain

FIG. 4 The U1 RNA binding surface of the U1 A protein. Main-chain trace (white) with some external residues of the highly conserved RNP1 and RNP2 segments (blue). The Lys and Arg residues are coloured according to the magnitude of the reduction in the U1 RNA binding affinity on replacement by Gln: Arg 52 (red), complete loss of specific binding; Lys 22, Lys 27, Lys 50 and Lys 80 (pink), up to 100-fold reduction; Lys 23, Arg 47, Arg 83 and Lys 88 (yellow), up to 10-fold reduction; Arg 7, Lys 20, Lys 28, Lys 60 and Lys 70, no effect on RNA binding (pale blue). Ser 46 and Ser 48 (green) may be involved in discrimination of U1 RNA and U2 RNA.



10 and 13 nucleotides, respectively, and larger loops are probably more flexible. The first six nucleotides of the U1 RNA loop II are identical to the corresponding part of the U2 RNA loop IV, and they probably interact with the region where U1 A and U2 B' proteins have the same residues, such as Arg 52, Gln 54, Phe 56, Tyr 13, Asn 15, Asn 16 and Asn 18 of RNP1 and RNP2. Arg 52, mutation of which completely abolishes U1 RNA binding, may form hydrogen bonds to a base. All these residues, except Phe 56, can form hydrogen bonds to bases; in the loop regions hydrogen-bond donors and acceptors of bases that might otherwise be used for base-pairing can also bind to amino acids. Phe 56 and Tyr 13 may intercalate between bases. C7 and/or C9 of the U1 RNA loop II probably interact with residues on the β 2 strand as discussed above. These considerations leave two possible ways of placing the U1 RNA loop II on the surface of the four-stranded β sheet. If the side of the loop that forms the continuation of the major groove interacts with the protein surface, the stem will point upwards on the β sheet in Fig. 4. If the other side of the loop is on the protein surface the stem will extend downwards. We favour the upwards model, because the loop can have extensive interactions with a pair of jaws formed between the two loops and the stem can sit on the sheet.

Crystallographic studies of the RNA-U1 A protein complex are underway and should provide insight into the RNA-protein interaction and the sequence specificities of the RNP domain that have crucial roles in RNA storage, transport and splicing.

In our crystal structure, the RNP domain forms a dimer related by a non-crystallographic dyad axis (Fig. 2b). Many hydrophobic amino acids are found in this interface (Fig. 3b): this suggests that the dimer is not an artefact of crystallization, but that the RNP domain can form such dimers either with itself or with the cognate domains in other proteins. The formation of the dimer leaves the RNA-binding surface of the domain free, and also the putative binding surface for the U2 A' protein on U2 B'. There are several possibilities for dimer formation including (1) heterodimer formation between the N- and C-terminal domains of U1 A protein itself; or (2) dimer formation between U1 A and U2 B', either pairing N-terminal domain with N-terminal domain and C-terminal domain with C-terminal domain, or pairing N-terminal domain of one protein with C-terminal domain of the other. One reason for such association could be the interaction between U1 snRNP and U2 snRNP to bring together the 5' exon-intron junction and the branch site²⁻⁴. □

Received 11 October; accepted 23 October 1990.

- Dreyfuss, G., Swanson, M. S. & Pinol-Roma, S. *Trends biochem. Sci.* **13**, 86-91 (1988).
- Maniatis, T. & Reed, R. *Nature* **325**, 673-678 (1987).
- Sharp, P. *Science* **235**, 766-771 (1987).
- Steitz, J. A. *Sci. Am.* **256** (June) 36-41 (1988).
- Bandziulis, R., Swanson, M. S. & Dreyfuss, G. *Genes Dev.* **3**, 431-437 (1989).
- Query, C. C., Bentley, R. C. & Keene, J. D. *Cell* **57**, 89-101 (1989).
- Adams, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K. & Dreyfuss, G. *Molec. cell. Biol.* **6**, 2932-2943 (1986).
- Sachs, A. B., Bond, W. M. & Kornberg, R. D. *Cell* **45**, 827-835 (1986).
- Mattaj, J. W. *Cell* **57**, 1-3 (1989).
- Theissen, H. *et al.* *EMBO J.* **5**, 3209-3217 (1986).
- Sillekens, P. T. G., Habets, W. J., Beijer, R. P. & van Venrooij, W. J. *EMBO J.* **6**, 3841-3848 (1987).
- Scherly, D. *et al.* *EMBO J.* **8**, 4163-4170 (1989).
- Scherly, D., Boelens, W., Dathan, N. A., van Venrooij, W. J. & Mattaj, J. W. *Nature* **345**, 502-506 (1990).
- Lutz-Freyermuth, C., Query, C. C. & Keene, J. D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6393-6397 (1990).
- Patten, J. R. & Pederson, T. *Proc. natn. Acad. Sci. U.S.A.* **85**, 747-751 (1988).
- Bach, M., Knol, A. & Lührmann, R. *Nucleic Acids Res.* **18**, 449-457 (1990).
- Scherly, D., Dathan, N. A., Boelens, W., van Venrooij, W. J. & Mattaj, J. W. *EMBO J.* **9**, 3675-3681 (1990).
- Habets, W. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 2421-2425 (1987).
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. *Meth. Enzym.* **185**, 60-89 (1990).
- Arndt, U. W. *Meth. Enzym.* **114**, 472-485 (1985).
- Sheldrick, G. M. in *Crystallographic Computing 3* (eds Sheldrick, G. M., Krüger, C. & Goddard, R.) 175-189 (Oxford University Press, Oxford, 1985).

- Wang, B. C. *Meth. Enzym.* **115**, 90-112 (1985).
- Leslie, A. G. W. in *Computational Aspects of Protein Crystal Analysis* (eds Helliwell, J. R., Machin, P. A. & Papiz, M. Z.) 35-90 (SERC, Daresbury Laboratory, Daresbury, UK, 1987).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr.* (in the press).
- Brünger, A. T. *J. molec. Biol.* **203**, 803-816 (1988).
- Hendrickson, W. A. & Konert, J. H. in *Structure, Conformation and Evolution* Vol. 1 (ed. Srinivasan, R.) 43-57 (Pergamon, Oxford, 1981).
- Sauadek, V. *et al.* *J. molec. Biol.* **207**, 405-415 (1989).
- Ghetti, A., Bolognesi, M., Cobiainchi, F. & Morandi, C. *Molec. Biol. Rep.* **14**, 87-88 (1990).
- Jones, T. A. in *Computational Crystallography* (ed. Sayre, D.) 303-317 (Clarendon, Oxford, 1982).
- Steitz, T. A. *Quart. Rev. Biophys.* **23**, 205-280 (1990).
- Morris, C. E., Klement, J. F. & McAllister, W. T. *Gene* **41**, 193-200 (1986).
- Robertus, J. D. *et al.* *Nature* **250**, 546-551 (1974).
- Kim, S. H. *et al.* *Science* **185**, 435-440 (1974).
- Moras, D. *et al.* *Nature* **288**, 669-674 (1980).
- Schevitz, R. W. *et al.* *Nature* **286**, 346-351 (1979).
- Rould, M. A., Perona, J. J., Söhl, D. & Steitz, T. A. *Science* **246**, 1135-1142 (1989).

ACKNOWLEDGEMENTS. We thank A. Klug, M. F. Perutz, W. Sundquist, D. Bogenhagen, A. Leslie, D. Rhodes and R. Turner for their comments and criticism, C. Pritchard for collaboration on other aspects of the RNA-protein interaction, D. Bogenhagen for advice on RNA band-shift experiment, T. A. Jones for program O and help in model building, A. Lesk for discussion and drawing figures, P. McLaughlin and B. Luisi for advice on protein crystallization, I. Mattaj for communicating unpublished results, C. Hellon for excellent workmanship, and J. Fogg and T. Smith for oligonucleotide synthesis. This work was supported by the MRC and the HFPS and NIH. T.H.J. thanks the Max-Planck-Gesellschaft for financial support (A. Butenandt fellowship). Crystal coordinates will be deposited in Brookhaven Protein Databank.

LETTERS TO NATURE

Evidence for long-term brightness changes of solar-type stars

Sallie Baliunas*‡ & Robert Jastrow†

*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

†Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

CHANGES in the brightness of the Sun may introduce further uncertainties into forecasts of global warming by the greenhouse effect. The Sun is known to vary in brightness, on a timescale of years, by 0.1% in phase with changes in magnetic activity during the solar cycle¹⁻³, and variations of up to 0.4%, also correlated with surface magnetic activity, have been found in stars similar to the Sun⁴. To delimit the magnitude of solar luminosity variations on a timescale of centuries, we have looked at the magnetic behaviour of a number of solar-type stars over several years. Observed in random phases of their long-term variability, they

give a sample of the behaviour of a solar-type star over a long period of time. We find indirect evidence that these stars undergo brightness changes of more than the 0.1% observed during the last solar cycle, a result that calls into question the assumption of a constant Sun in calculations using general circulation models for climate forecasting.

We have initiated a programme aimed at determining solar changes over longer periods of time using the hypothesis of stellar equivalence (that all stars of the same age and mass have nearly identical properties with respect to magnetic activity). In the first phase of the programme we have analysed changes in magnetic activity in solar-type stars, using twenty years of calcium H-line and K-line fluxes obtained with Mount Wilson Observatory's 60-inch and 100-inch reflectors. The H and K fluxes are a measure of stellar magnetic activity⁵⁻⁸. We calculated the distribution of magnetic activity for 74 solar-type stars (main-sequence stars with masses and ages similar to the Sun's).

The stars in our sample are being used in two of the programmes of the Mount Wilson Observatory HK project⁹. Our criteria for the selection of the main-sequence stars are: (1) mass such that the blue-visual (B-V) magnitude is in the range 0.60-0.76; (2) age similar to the Sun. The range of (B-V) corresponds to ~1.10-0.95 $M_{\odot}$ (ref. 10), which is narrow enough to

‡ Also at: Center of Excellence in Information Systems, Tennessee State University, and Department of Physics and Astronomy, Dartmouth College.

focus on solar-mass stars, but generous enough to gather sufficient stars for statistical purposes.

Determination of the ages of the individual field dwarf stars is crucial to our analysis for the reasons indicated below, but subject to substantial uncertainties. For a preliminary age determination, we used the calibration of age and magnetic activity for main-sequence stars, which indicates that activity decreases with age, and selected stars with average levels of magnetic activity relatively close to that of the Sun. This procedure is qualitatively valid; that is, it will tell us whether a star is old or young compared to the Sun. The range of magnetic activity that we chose to correspond to solar age is $0.13 < \langle S \rangle < 0.20$, where $\langle S \rangle$ is the time-averaged value of the relative calcium H and K flux measured by the Mount Wilson instrument, defined as the flux in narrow passbands centred on the calcium H and K emission cores compared in a ratio to the flux in the nearby stellar continuum^{8,11}. The value of S for the Sun ranges between 0.164 and 0.178 during the 11-yr sunspot cycle (cycle 20) and averages ~ 0.171 (ref. 8). According to the calibration, the selected range in S corresponds to ages of at least 3×10^9 years¹², and to an upper limit not greater than $9\text{--}10 \times 10^9$ years.

Of our 74 solar-type stars, thirteen have been studied monthly since 1966⁸, and nightly since 1980⁹. There are sporadic measurements beginning in 1978 for the other 61 stars¹³ and we have included these to improve the statistics. We calculated averages at 0.1-yr intervals for each star and treated each average as a snapshot of one brief interval in the long-term behaviour of a solar-type star. Concentrating on the 13 stars that make up the bulk of the data, we note that the time series fall into two groups. Four stars are magnetically 'flat', showing little or no variation of magnetic activity with time, and nine stars show a range of magnetic activity similar to the solar cycle. The magnetic activity of the four 'flat' stars are almost always lower than the activity of the 'cyclic' stars.

Examples of 23-yr time series for solar-type stars are shown in Fig. 1. The star HD81809 (upper panel) is an example of stars having cyclic fluctuations similar to the sunspot cycle. The star HD9562 is an example of the group of stars that are magnetically 'flat'.

Figure 2 gives the frequency distribution of magnetic activity for the 74-star subsample. The histogram shows at a glance how much time, in a statistical sense, the solar-type stars in the sample spend at each level of magnetic activity.

The broad peak centred at $S \approx 0.17$ corresponds to the distribution expected for a sample of stars varying periodically, observed at random phases in their activity cycles. This distribution has approximately the mean level expected if the average S during the cycles in these stars matched the mean S of the Sun's cycle, namely, 0.171. The width of the distribution is comparable to but somewhat broader than the Sun's range in S over a typical solar cycle.

There is a narrow peak at low levels of magnetic activity that is apparently separated from the broad distribution. Its centre is at $S \approx 0.145$, substantially lower than the values of $S \approx 0.164$ at the bottom of the sunspot cycle and only slightly above the value $S \approx 0.14$ which corresponds to near-zero surface magnetism^{14,15}. We suggest that this narrow peak at low values of S indicates that many solar-type stars have long-term lulls in activity, analogous to the Maunder minimum of the Sun, with magnetic flux values much lower than observed at the minimum of the Sun's 11-yr cycle.

The term Maunder minimum is usually used in the specific context of the seventeenth-century observations of the Sun. Here we broaden its use to solar-type stars exhibiting prolonged low levels of magnetic activity. This generalized definition is based on the fact that the radiocarbon record identifies a period of low solar magnetic activity in the seventeenth century that coincided with the virtual disappearance of sunspots during the Maunder minimum¹⁶. Further, the radiocarbon record shows that the Sun has exhibited such periods of low magnetic activity

every few hundred years for the past several millennia¹⁷. We conclude that the reduction of sunspots observed in the seventeenth century may be a general characteristic of solar-type stars that is best described by a low mean level of magnetic activity, rather than the near-disappearance of sunspots. The varying level of magnetic activity is a more quantitative indicator of changes in the Sun's physical state than the historical record of sunspots.

We also assume that Maunder minima are characterized by magnetic 'flatness', because the Sun was magnetically flat during its Maunder minimum—that is, not only were there few visible signs of magnetic activity, in the form of sunspots and aurorae, but the solar cycle was also not evident¹⁶. Because magnetic flatness cannot be deduced from the radiocarbon record alone, it is not clear whether flatness is common to all or most of the periods of low magnetic activity. Here, however, we assume that magnetic flatness is also a defining property of Maunder minima.

In Fig. 2 the gap between the broad peak centred on $S \approx 0.17$ and the narrow peak at low values of S indicates that the frequency distribution of magnetic activity may be bimodal. The bimodality of the distribution is not essential to our result, which is that some solar-type stars show behaviour similar to that of the Sun in the Maunder minimum, with very low levels of magnetic activity. But if the gap is real, it indicates that solar-type stars spend some of their time in cyclic states of normal magnetic activity, some time in the low and flat Maunder minimum states, and not much time in transit between the two states.

The interpretation of the low- S stars in terms of a Maunder minimum is tentative because they may be old, and therefore in a permanent, not temporary, state of low magnetic activity. But as indicated above, a low level of magnetic activity by itself is not a precise indicator of age. For example, if we had measured a low level of magnetic activity for the Sun at the time of its Maunder minimum, we would have been misled into inferring a spuriously great age. Other evidence that magnetic activity is not a good measure of age in solar-type stars comes from a cyclic active star in the sample, HD103095 (Groombridge 1830), which has a large-amplitude magnetic cycle, similar to the Sun's⁹. This star is a high-velocity subdwarf with low metallicity¹⁸, and is clearly an old star. Its behaviour counters the

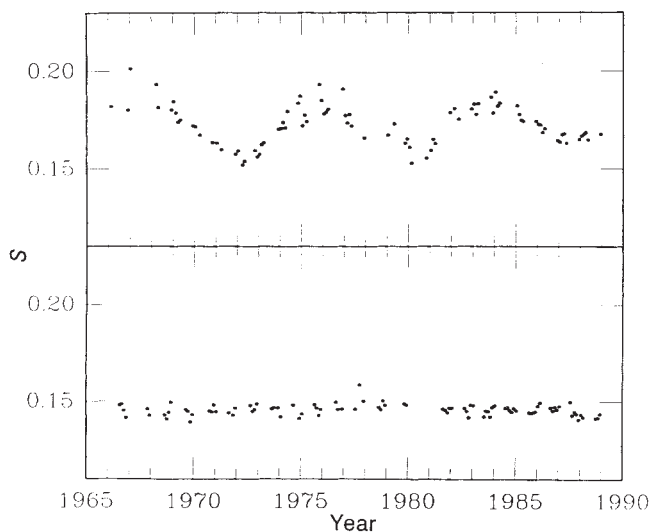


FIG. 1 Records of long-term fluctuations in magnetic activity, represented by S in two solar-type stars. Increased values of S signal higher levels of magnetic activity. Observations were made on a monthly basis before 1980, and on a nightly basis thereafter. Plotted points after 1980 are monthly averages. The star HD81809 (upper panel) has an amplitude of variation, a mean value and cycle period (~ 9 yr) that are close to those of the Sun. The star HD9562 (lower panel) has S values that are constant with time.

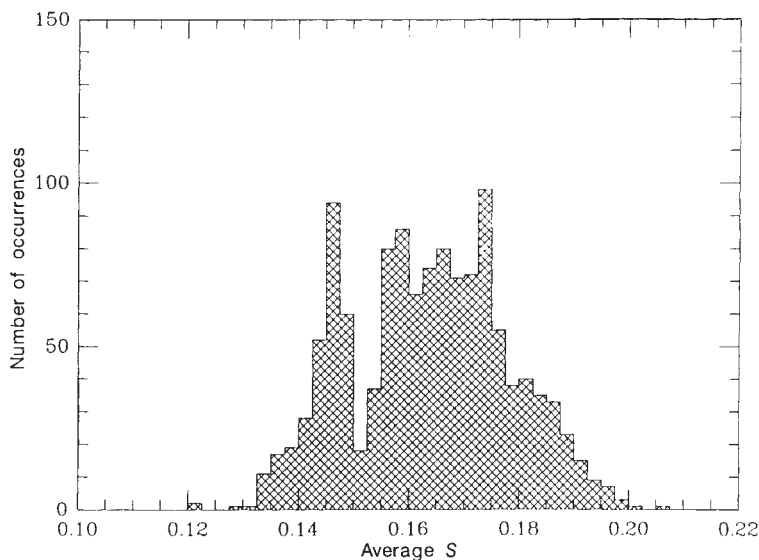


FIG. 2 The frequency distribution of magnetic activity in 74 solar-type stars. The measurements have been averaged over 0.1-yr intervals. The broad peak at higher values of magnetic activity ($S > 0.15$) corresponds to the broad distribution expected for stars varying periodically and observed at random phases in their cycles. The narrow peak at low values of S (< 0.15) may represent stars sampled in Maunder minima.

notion that extremely old stars must have low and non-variable magnetic activity.

Further support for the hypothesis that the low level of magnetic activity in the solar-type stars represents a Maunder minimum, and not advanced stellar age, comes from a detailed examination of the individual ages of the 13 stars for which there is nearly continuous sampling coverage over 23 years. Four stars out of the thirteen have low and essentially flat values of S . The low values contribute to the existence of the narrow low- S peak in the distribution. For two of these stars (HD9562 and HD143761), independent values of the ages can be determined from rotation (inferred from Doppler broadening^{19,20}) and turn out to be the Sun's age or younger. Thus, age cannot be the explanation for the magnetically flat records and low fluxes of these stars. Because their magnetic activity levels are extremely low and they are also magnetically flat, they satisfy the criteria for a stellar Maunder minimum. For the other two magnetically flat stars, HD3795 and HD43587, no indication of age is available, other than their low average level of magnetic activity and they therefore neither confirm nor deny the Maunder minimum interpretation.

The finding that roughly a third (four) of solar-type stars seem to be in Maunder minima at any one time is in good agreement with the ¹⁴C record of solar magnetic activity, which suggests that the Sun has spent about one-third of its time in the past several millennia in magnetic minima¹⁷.

It is significant that the four magnetically flat stars have low levels of magnetic activity compared to the level at the minimum in the solar cycle. The mean values of S for HD3795, HD9562, HD43587 and HD143761 are 0.158, 0.146, 0.158 and 0.147, respectively, only slightly lower than the Sun's value of S at the lowest point of the solar cycle^{8,21}, 0.164. (Two stars are completely in the lower distribution, and two are either in the gap or in the lowest bins of the upper distribution.) A large portion of the value of S , for these stars and for the Sun, is from non-magnetic contributions at and near the surface of the star, however, which include contamination by photospheric light counted within the passbands of our instrument²² as well as non-magnetic heating¹⁵. These non-magnetic sources account for ~0.14 of the measured S (ref. 14). Thus the value of S for the Sun at the lowest point of its cycle is 0.024 above the extrapolated level for zero magnetism. The average values of S for the four stars at Maunder minimum are considerably lower, ranging from 0.006 to 0.018 above the value corresponding to near-zero magnetism. Assuming a roughly linear dependence of average surface magnetism on H and K flux, the surface magnetism of these four stars is between one-quarter and three-quarters

of the surface magnetism of the Sun at the lowest point of its sunspot cycle.

We suggest the fact that the minimum of the solar cycle and the Maunder minimum have a similar appearance (that is, in both cases the sunspot number is near zero) has led to incorrect conclusions about the properties of the Sun in a Maunder minimum. Our analysis of the observations indicates that the level of magnetic activity is much lower in a Maunder minimum than in a sunspot minimum. Eddy has reached a similar conclusion on the basis of the radiocarbon record¹⁶.

Magnetic flux and brightness changes are positively correlated in the sunspot cycle¹⁻³, and also in the activity cycles that occur on solar-type stars⁴. If we assume that a positive correlation (not necessarily the same quantitative relationship) also holds for the Maunder minimum, our results suggest that when a solar-type star enters or leaves a Maunder minimum, its brightness may change by more than the 0.1% observed during the Sun's 11-yr cycle. A change of only 0.22–0.55% has been estimated to be sufficient to account for the 0.4–0.6 °C global mean temperature change in the Little Ice Age²³. Our results, which suggest the possibility of changes of several tenths of a per cent, may have significant implications for climate change over century-long timescales²³⁻²⁶. □

Received 11 June; accepted 19 October 1990.

- Willson, R. C., Hudson, H. S., Fröhlich, C. & Brusa, R. W. *Science* **234**, 1114–1117 (1986).
- Willson, R. C. & Hudson, H. S. *Nature* **332**, 810–812 (1988).
- Hickey, J. R., Alton, B. M., Kyle, H. L. & Hoyt, D. *Space Sci. Rev.* **48**, 321–342 (1988).
- Radick, R. R., Lockwood, G. W. & Baliunas, S. L. *Science* **247**, 39–44 (1990).
- Eberhard, G. & Schwarzschild, K. *Astrophys. J.* **38**, 292–295 (1912).
- Babcock, H. W. & Babcock, H. D. *Astrophys. J.* **121**, 349–366 (1955).
- Skumanich, A., Smythe, C. & Frazier, E. N. *Astrophys. J.* **200**, 747–763 (1975).
- Wilson, O. C. *Astrophys. J.* **226**, 379–386 (1978).
- Baliunas, S. L. & Vaughan, A. H. *A. Rev. Astr. Astrophys.* **23**, 379–412 (1985).
- VandenBerg, D. A. & Bridges, T. J. *Astrophys. J.* **278**, 679–688 (1984).
- Vaughan, A. H., Preston, G. W. & Wilson, O. C. *Publ. astr. Soc. Pac.* **90**, 267–274 (1978).
- Soderblom, D. R., Duncan, D. K. & Johnson, D. R. H. *Astrophys. J.* (in the press).
- Vaughan, A. H. & Preston, G. W. *Publ. astr. Soc. Pac.* **92**, 385–391 (1980).
- Schrijver, C. J., Cote, J., Zwaan, C. & Saar, S. H. *Astrophys. J.* **337**, 964–976 (1989).
- Schrijver, C. J., Dörsen, A. K. & Radick, R. R. *Astrophys. J.* **341**, 1035–1044 (1989).
- Eddy, J. A. in *Secular Solar and Geomagnetic Variations in the Last 10,000 Years* (eds Stephenson, F. R. & Wolfendale, A. W.) 1–24 (Kluwer, Dordrecht, 1988).
- Damon, P. E. in *The Solar Output and Its Variation* (ed. R. White) 429–448 (Colorado Ass. Univ. Press, Boulder, 1977).
- Tomkin, J. & Bell, R. A. *Mon. Not. R. astr. Soc.* **163**, 117–127 (1973).
- Pallavicini, R., Cerruti-Sola, M. & Duncan, D. K. *Astr. Astrophys.* **174**, 116–126 (1987).
- Soderblom, D. R. *Astrophys. J. Suppl.* **53**, 1–15 (1983).
- White, O. R., Rottman, G. J. & Livingston, W. C. *Geophys. Res. Lett.* **17**, 575–578 (1990).
- Hartmann, L., Baliunas, S. L., Duncan, D. K. & Noyes, R. W. *Astrophys. J.* **279**, 778–784 (1984).
- Wigley, T. M. L. & Kelly, P. M. *Phil. Trans. R. Soc. A* **330**, 547–560 (1990).
- Hansen, J. et al. *Science* **213**, 957–966 (1981).
- Reid, G. C. & Gage, K. S. in *Secular Solar and Geomagnetic Variations in the Last 10,000 Years* (eds Stephenson, F. R. & Wolfendale, A. W.) 225–243 (Kluwer, Dordrecht, 1988).
- Gilliland, R. L. *Clim. Change* **4**, 111–131 (1982).

ACKNOWLEDGEMENTS. We thank our former and current colleagues of the Mount Wilson Observatory and the HK Project: W. Bennett, J. Briggs, R. Donahue, D. Duncan, J. Frazer, J. Horne, H. Lanning, A. Misch, J. Mueller, D. Poppe, A. Porter, C. Robinson, C. Shelton, T. Soyumer, R. Ulrich, A. Vaughan, L. Webster, O. Wilson and L. Woodard. We also thank P. Foukal and G. Withbroe for criticisms. This research was supported by the NSF, the Langley-Abbot Program, the Scholarly Studies Program, and other funds of the Smithsonian Institution, the Center of Excellence in Information Systems of Tennessee State University, the Mobil Foundation Inc. and the Lounsberry Foundation.

Modelling viscous segregation in immiscible fluids using lattice-gas automata

H. W. Stockman*, C. T. Stockman† & C. R. Carrigan‡

* Division 6233, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

† Science Applications International Corporation, Albuquerque, New Mexico 87106, USA

‡ Earth Sciences Division, L-206, Lawrence Livermore Laboratory, Livermore, California 94550, USA

WHEN immiscible liquids with different viscosities are forced to flow through a channel, the more viscous liquid tends to concentrate in the centre. This process influences the flow of oil–water mixtures in pipelines^{1,2}, the extrusion of polymers³, the flow of blood in small arteries⁴ and of magmas approaching the Earth's surface during a volcanic eruption⁵. Segregation occurs because it minimizes the pressure required to maintain a given flow rate^{6,7}; there is still scant information, however, on the timescales and fluid configurations involved in the approach to the equilibrium state. Explicit solutions of the time-dependent Navier–Stokes equation using numerical (for example, finite-element³) methods are possible only for simple interface shapes. Here we use an alternative approach to the study of viscous segregation, involving immiscible lattice-gas automata. Two-dimensional calculations exhibit the expected segregation for a variety of starting configurations, and for an initially homogeneous emulsion the fluid must flow 50–100 times the channel width before most of the more viscous fluid reaches the channel centre. For constant flow rates, segregation occurs so as to progressively decrease the pressure.

Lattice-gas automata are collections of discrete particles constrained to move on fixed geometric lattices^{8,9}. At each time step in a calculation, the particles translate along the links of the lattice, undergoing simple collisions that conserve momentum and particle number. After a simple scaling, macroscopic averages of particle velocity over time and space obey the Navier–Stokes equations for incompressible flow. The dynamic viscosity μ is a function of the allowed collisions; the greater the number of possible collisions, the lower the viscosity. This study uses a two-dimensional hexagonal lattice, and the particles have unit speed and six possible velocities.

Rothman and Keller¹⁰ introduced immiscible lattice-gas automata (ILGA). ILGA have two types of particles, 'white' and 'black'; collision rules are designed to cause particles to be attracted to like-coloured particles, subject to the constraint of momentum conservation. We denote the two fluids with the subscripts b (black) and w (white) and assume that the black fluid is the more viscous. ILGA correctly model surface tension and preserve the most important aspects of hydrodynamic flow.

In our variant of ILGA, the fluids are contained in a channel with horizontal 'no-slip' walls; the channel is typically 200 lattice units long and $104(\sqrt{3})/2$ or $200(\sqrt{3})/2$ lattice units wide and the density ρ is 2.3–3.0 particles per site. The walls are equally wetted by both fluids. A pressure gradient is applied by increasing the horizontal momenta of particles selected at random, resulting in a net flow from left to right; the pressure is applied uniformly across the width of the channel. The x direction is parallel to the channel axis, and the boundaries wrap, so that

particles leaving the right end of the channel at (x, y) reenter on the left side at $(-x, y)$. Two methods are used to vary μ : in the first method, different collision rules are used for sites that predominantly contain white particles and for sites that predominantly contain black particles; in the second, the collision step is performed less frequently for sites that predominantly contain black particles. Both methods give substantially identical results, but the second allows a greater range of μ_b/μ_w . Surface tension σ is a function of ρ , and $\sigma \approx 0$ for $\rho < 2.1$. The two fluids have the same density, and there are no gravitational effects. We ran our calculations on IBM-compatible microcomputers, and typically updated sites at a rate of 5×10^4 – 5×10^5 s⁻¹.

For comparison with other studies, it is useful to characterize the flow with two more dimensionless parameters. We define the capillary number $Ca_b = U\mu_b/\sigma$, and the Reynolds number $Re_b = UL/\nu_b$, where U is the average speed in the x direction, L is the width of the channel and ν is the kinematic viscosity ($\nu = \mu/\rho$); the parameter J^* of Preziosi *et al.*¹¹ is $\sim Re_b/2Ca_b$. Ca_b is a rough indicator of the ratio of viscous to surface-tension forces. If Ca_b is too small, the liquids minimize surface tension by segregating along opposite walls of the channel with one interface roughly parallel to the flow. Unfortunately, the discrete nature of the ILGA limits us to modest values of Ca_b ; U must be kept low (≤ 0.2 lattice units per time step) to avoid high-order corrections to the Navier–Stokes equation^{12,13}, and the surface tension must be kept finite, or the white and black particles rapidly interdiffuse. By using narrower channels (thus increasing the shear rate), segregation can be induced at $U \approx 0.1$, but mean-free-path effects become more significant. With these restrictions we are able to vary Re_b from 0 to 40, μ_b/μ_w from 1.6 to 10, and Ca_b from 0 to 10. For comparison, the flow of SAE 30 oil and water in a 15-cm diameter pipe at 15 cm s⁻¹ and 25 °C would be characterized by $Ca_b \approx 1.3$, $Re_b \approx 56$ and $\mu_b/\mu_w \approx 400$. For mixtures of silicic magmas, $Re \approx 1$ and μ_b/μ_w is 2–200 (ref. 5) (Ca_b is effectively infinite, because most magmas are miscible). With magmas (and polymers), chemical diffusion is so slow that the two fluids are effectively immiscible on the timescale of emplacement, but with vanishingly small σ .

In total, we have run over 50 calculations; 4 representative runs are shown in Fig. 1, each column consisting of 7 'snapshots' from a single run. For Fig. 1a, the initial configuration is a homogeneous emulsion. By step 31,500, the liquids have segregated into a few slugs; because the slugs are relatively large and the surface tension is low, they tend to be crescent-shaped or elongate, rather than rounded. By step 51,500, the black liquid is clearly concentrated in the centre of the channel; at this point the fluid has travelled a horizontal distance equivalent to about 60 channel widths. It takes another 2×10^5 steps, however, to complete the segregation.

Figure 1b, c begin with the viscous black fluid distributed evenly along the channel walls, or against one wall, respectively. Both systems eventually redistribute to form viscous-centred flows, and waves play an important part in the redistribution. In Fig. 1b, viscous black waves fold over and trap white fluid within the viscous walls, and the tips of black waves break off and drift toward the centre of the channel, coalescing in the centre of the flow. In Fig. 1c, tips of white waves break off and become entrapped within the viscous layer; the small blobs of trapped white fluid eventually coalesce and form a continuous, lubricating band near the bottom wall; once this band forms, the shear rate in the main viscous layer is lessened, the wave action decreases, and the segregation slows down dramatically. The 'wrap' boundary condition used in our calculations artificially stabilizes the initial configurations with respect to long-wavelength disturbances; we have run simulations with length/width varying by a factor of three, however, and the same mechanism of wave formation seems operative across this range.

Like Fig. 1a, Fig. 1d begins with a homogeneous emulsion but the surface tension in Fig. 1d is much lower than in Fig. 1a

and the shear rate is higher. Therefore, even in the final 'equilibrium' configuration, packets of fluid are constantly ripped off the sides of the central flow. These packets can travel horizontal distances equal to several channel widths before reattaching to the central viscous flow.

In Fig. 2, the horizontal component of velocity for the final configuration of Fig. 1a is compared with an analytical solution. The analytical solution assumes flat interfaces, and is derived by requiring continuity of shear and velocity at the black/white interfaces¹. The ILGA velocities are ~15% lower than the analytical solution in the centre of the flow. The relatively high flow rate in the centre of the channel, and the consequent error in apparent μ ^{12,13}, can account for a 2–4% difference between the ILGA calculation and the theoretical curve (we base this error estimate on 46 Poiseuille-flow calibrations run at $U = 0.025$ – 0.25). Most of the discrepancy arises from the simple interface shape assumed in the analytical solution. With real multiviscosity fluids and with the ILGA, the interfaces tend to ripple, disrupting the flow and consuming energy in opposition to surface tension. As a consequence of the rippling effect, viscous-centred flows are virtually always slower (for a given

pressure gradient) than the predictions of models that use smooth interfaces¹⁴.

Figure 3 shows the drop in pressure for the calculation in Fig. 1a, as a function of time. The ILGA has clearly proceeded along a path of lowering pressure, in accordance with the viscous-dissipation principle¹.

Three of the potential limitations of ILGA calculations are not significant here. The first is galilean invariance¹⁵, which can be illustrated with the velocity equation for a monocoloured lattice gas

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{\nabla P}{\rho} - g(\rho)(\mathbf{u} \cdot \nabla)\mathbf{u} + \nu \nabla^2 \mathbf{u}$$

where $\mathbf{u}$ is the macroscopic average of particle velocity, t is time, P is pressure and $g(\rho) = (3 - \rho)/(6 - \rho)$. In the Navier-Stokes equation, unless $g(\rho) = 1$ the advection of vorticity does not obey the galilean transforms; however, for our lattice gases, $g(\rho) \leq \frac{1}{2}$. For calculations that involve only flow, the lattice-gas equation can be made galilean-invariant by defining a new velocity $\mathbf{u}^* = g(\rho)\mathbf{u}$, and $P^* = g(\rho)P$. Unfortunately, tracer particles or blobs of immiscible liquid advect with the particle

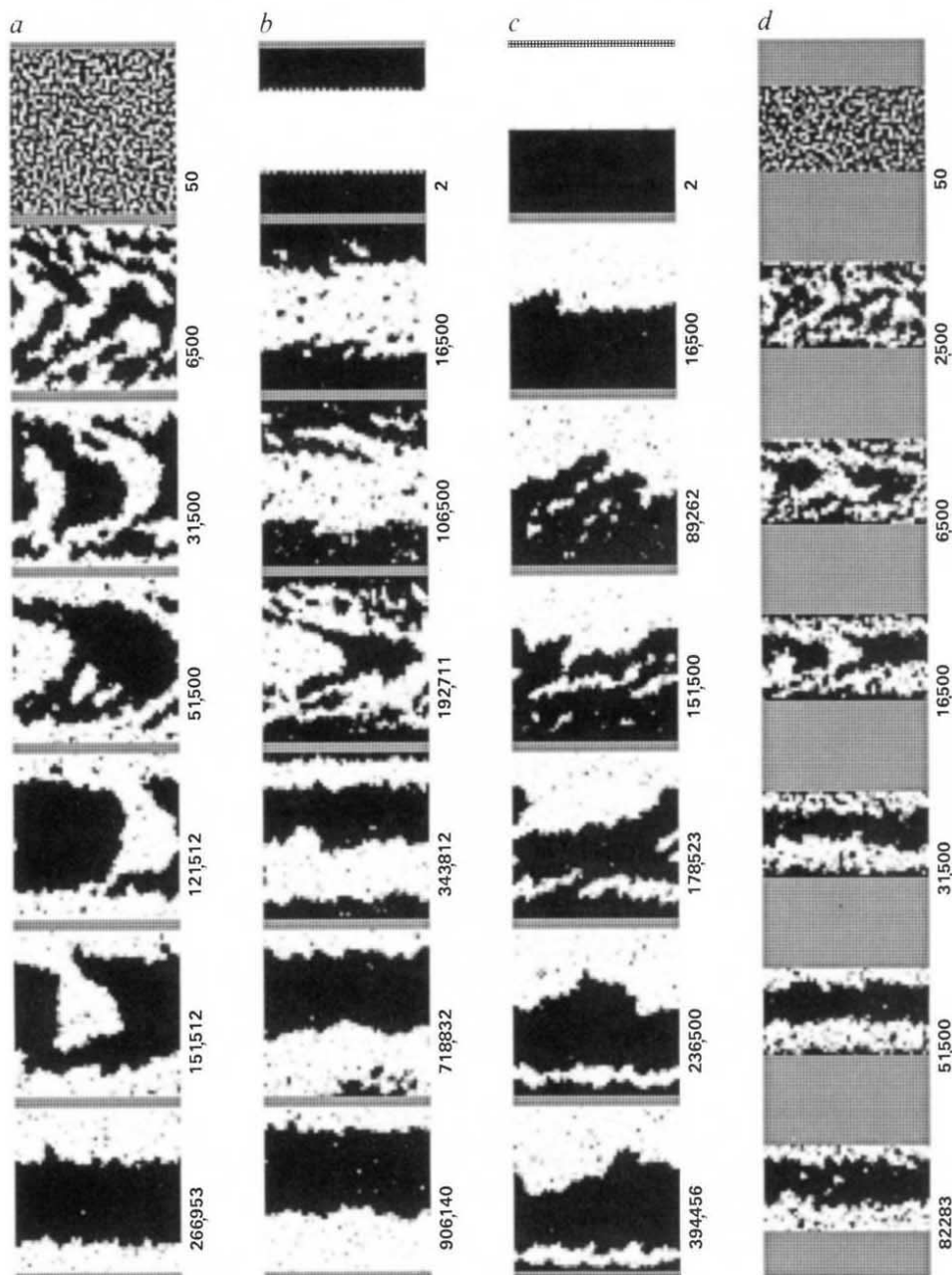


FIG. 1 Segregation of viscous (black) and less viscous (white) fluids; flow is from left to right. Each column is a separate run from a different set of initial conditions, and the time step is given next to each frame. Concentrations are averaged over sets of 16 sites, and solid walls are denoted by the square mesh pattern. The pressure was continuously adjusted to maintain $U = 0.200 \pm 0.002$ lattice units per time step per particle. In all four runs, the numbers of black and white particles were equal, and $Ca_0 \approx 5$. *a*–*c* were run with $\mu_b/\mu_w = 8$, $\rho = 2.5$, $Re_b = 5$ and $L = 200(\sqrt{3})/2$ lattice units; only the initial configurations were varied. *d* was run with $\mu_b/\mu_w = 2.5$, $\rho = 2.3$, $Re_b = 12$ and $L = 104(\sqrt{3})/2$.

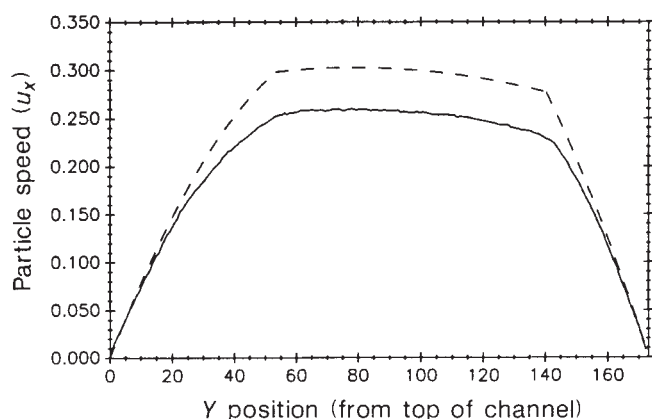


FIG. 2 Average particle speeds in the x (horizontal) direction for Fig. 1a, in the final configuration. The continuous curve is the ILGA calculation and the dashed curve is the analytical solution¹. For the analytical solution, the black/white interfaces are assumed to be perfectly flat and horizontal and are placed at the average position observed in the final frame of Fig. 1a. The interfaces of the ILGA are rippled, much like a real fluid, which lowers the flow rate relative to the analytical solution.

velocity $\mathbf{u}$, not $\mathbf{u}^*$, and the flow and concentration equations become decoupled if $\mathbf{u}^*$ is used for the flow velocity. Galilean invariance is not crucial for this study, primarily because the $(\mathbf{u} \cdot \nabla)\mathbf{u}$ term tends to be small relative to $\nu \nabla^2 \mathbf{u}$; that is, the flow is dominated by viscous forces rather than inertial forces. We have run ILGA calculations with $\rho = 3$, so that $g(\rho)(\mathbf{u} \cdot \nabla)\mathbf{u} = 0$, and viscous-centred flows still develop in approximately the same number of time steps; hence, the segregation does not require inertial forces (we note that the principle of minimum entropy production¹⁶, which is the basis for the viscous-dissipation principle, implicitly assumes that inertial forces are negligible). We therefore report $\mathbf{u}$ rather than $\mathbf{u}^*$, and in calculating Re_b we absorb $g(\rho)$ in the viscosity.

Second, ILGA calculations can also suffer from 'pseudoforces', with magnitude proportional to $\rho g(\rho) \partial(u^2)/\partial y$, which can accelerate mass in a direction roughly perpendicular to the channel axis¹⁷, producing slight density variations that superficially resemble sound waves. Pseudoforces seem unlikely to be the cause of viscous segregation, however, because we still observe segregation for calculations with $g(\rho) = 0$.

Finally, the mean free path λ of ILGA particles may limit the accuracy of the simulations. For the white fluid, $\lambda \approx 3$ –4, and for the black fluid, $\lambda \approx 5$ –10 lattice units (we refer to the mean free path in an infinite monocoloured fluid; at the interface between fluids, particles are almost invariably deflected, so the

actual mean free path is typically smaller than λ). Consider a droplet of fluid with radius R ; if $R \leq \lambda$, the viscosity of the droplet will not be modelled accurately. If a droplet is small enough, its shape will be dominated by surface tension, however, and the viscosity of the droplet will be relatively unimportant. We approximate the radius $R_{s>\nu}$, below which surface tension dominates viscous forces, from the requirement that the pressure difference between the inside and outside of the droplet, $P_{in} - P_{out}$, exceeds the average pressure drop across the droplet, $2R_{s>\nu} \partial P / \partial x$. From LaPlace's formula for a two-dimensional bubble, $P_{in} - P_{out} = \sigma / R_{s>\nu}$, which yields $R_{s>\nu} \leq (\sigma / (2 \partial P / \partial x))^{1/2} \approx 10$ for the conditions of this study. Because $R_{s>\nu} \geq \lambda$, we can probably ignore any inaccuracy in the viscosity of small droplets.

The present ILGA technique is restricted to a relatively narrow range of parameter space; particle velocity cannot be made arbitrarily high, nor can σ be made arbitrarily low, which confines us to modelling flows where viscous forces barely dominate surface tension. Rothman and Gunstensen (personal communication) are developing ILGA with very small σ and three-dimensional galilean-invariant ILGA should soon be available^{18–20}, which will greatly expand the usefulness of the technique. □

Received 23 July; accepted 23 October 1990.

1. Joseph, D. D., Nguyen, K. & Beavers, G. S. *J. Fluid Mech.* **141**, 319–345 (1984).
2. Charles, M. E. & Lilleleht, L. U. *Can. J. chem. Engng* **43**, 110–116 (1965).
3. Karagiannis, A., Mavridis, H., Hrymak, A. N. & Vlachopoulos, J. *Polymer Engng Sci.* **28**, 982–988 (1988).
4. Quemada, D. in *Arteries and Arterial Blood Flow* (ed. Rodkiewicz, C. M.) 91–105 (Springer-Verlag, New York, 1983).
5. Carrigan, C. R. & Eichelberger, J. C. *Nature* **343**, 248–251 (1990).
6. Joseph, D. D., Renardy, M. & Renardy, Y. *J. Fluid Mech.* **141**, 309–317 (1984).
7. Than, P. T., Rosso, F. & Joseph, D. D. *Int. J. Engng Sci.* **25**, 189–204 (1987).
8. Frisch, U., Hasslacher, B. & Pomeau, Y. *Phys. Rev. Lett.* **56**, 1505–1508 (1986).
9. Frisch, U. et al. *Complex Systems* **1**, 648–707 (1987).
10. Rothman, D. H. & Keller, J. M. *J. stat. Phys.* **52**, 1119–1127 (1988).
11. Preziosi, L., Chen, K. & Joseph, D. D. *J. Fluid Mech.* **201**, 323–356 (1989).
12. Lim, H. A. *Phys. Rev. A* **40**, 968–980 (1989).
13. Hayot, F. *Complex Systems* **1**, 753–761 (1987).
14. Oilemans, R. V. A. & Ooms, G. in *Multiphase Science and Technology Vol. 2* (eds Hewitt, G. F., Delhay, J. M. & Zuber, N.) 427–472 (Hemisphere, Washington, DC, 1986).
15. d'Humières, D., Lallemand, P. & Searby, G. *Complex Systems* **1**, 633–647 (1987).
16. Schechter, R. S. *The Variational Method in Engineering*, 135–148 (McGraw-Hill, New York, 1967).
17. Dahlburg, J. P., Montgomery, D. & Doolen, G. D. *Phys. Rev. A* **36**, 2471–2474 (1987).
18. Gunstensen, A. K. & Rothman, D. H. *Physica D* (in the press).
19. Chen, S., Chen, H. & Doolen, G. D. *Complex Systems* **3**, 243–251 (1989).
20. Somers, J. A. & Rem, P. C. *Cellular Automata and Modeling of Complex Physical Systems* (eds Manneville, P., Boccaro, N., Vichniac, G. Y. & Bidaux, R.) 161–177 (Springer, Berlin, 1990).

Resolution of oxygen atoms in staurolite by three-dimensional transmission electron microscopy

Kenneth H. Downing*, Hu Meisheng†, Hans-Rudolf Wenk† & Michael A. O'Keefe‡

* Donner Laboratory, Lawrence Berkeley Laboratory,

† Department of Geology and Geophysics, University of California, and

‡ National Center for Electron Microscopy, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA

WITH advances in microscope design, high-resolution electron microscopy has become routine, and resolutions of less than 2 Å have been obtained for many inorganic crystals^{1–4}. The images provide a representation of the Coulomb potential (essentially the electron density), but interpretation generally requires comparing experimental images with calculations⁵. Because the images are two-dimensional representations of the full three-dimensional structure, information is invariably lost. In particular, oxygen atoms are normally not seen. The technique of electron crystallography, in which information from several views of a crystal is combined, has been developed to obtain three-dimensional

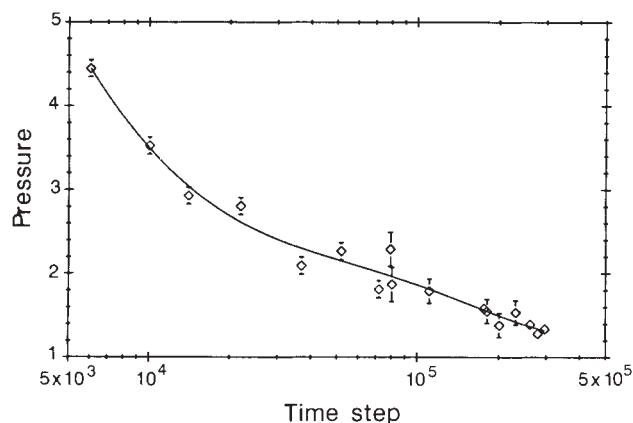


FIG. 3 The evolution of pressure with time in Fig. 1a; the ILGA apparently evolves in accordance with the viscous-dissipation principle (units of 'pressure' are the number of x -velocity perturbations required per time step to maintain an average flow rate of 0.2 lattice units per time step).

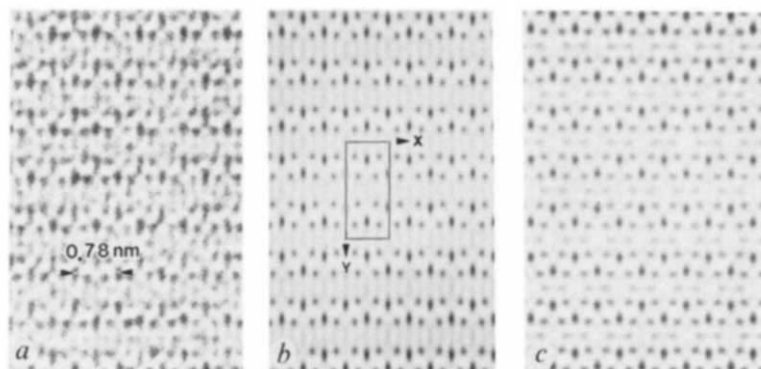


FIG. 1 Images of the structure of staurolite viewed along [001], JEOL ARM-1000 operating at 800 kV, Scherzer focus. *a*, Experimental image in thin area (< 40 Å). *b*, Image reconstructed using structure factors from the calculated Fourier transform of *a*. *c*, Multibeam dynamic contrast calculation¹² assuming the structure of staurolite, microscope conditions as described above and specimen thickness of 30 Å.

information on proteins⁶⁻⁸. Here we use this technique to obtain a three-dimensional reconstruction of the crystal structure of staurolite, a silicate mineral. We are able to identify the positions of oxygen atoms in the lattice. These results show the potential of the method for high-resolution structure determination on samples that are too small for the use of X-ray techniques.

Electron crystallography is in many ways analogous to X-ray crystallography, except that structure-factor phases are determined directly from high-resolution images. Because proteins are very susceptible to beam damage, most applications have provided resolution in the 10–20 Å range which allows morphological domains to be distinguished. The first atomic model of a protein to be determined by electron crystallography was recently obtained for the membrane protein bacteriorhodopsin⁹. From the electron density map at 3.5-Å resolution it was possible to trace the path of the polypeptide chain and define the location of amino acid side chains, although individual atoms could not be resolved. This same method should be applicable to other systems that are more stable in the electron beam and for which higher resolution can be obtained, provided that the effects of dynamical scattering can be minimized¹⁰.

We have used this method to determine the three-dimensional structure of staurolite to a resolution of 1.6 Å. Staurolite, a silicate mineral for which the complex structure is well known¹¹, is typical of many oxide structures with close-packed oxygens surrounding different types of cations in tetrahedral and octahedral coordination.

Crystals of staurolite $\text{HFe}_2\text{Al}_9\text{Si}_4\text{O}_{24}$ from Alpe Sponda, Switzerland, were sectioned perpendicular to the three main

directions [100], [010] and [001], and electron-transparent foils were prepared by ion-beam thinning. The space group is probably $C2/m$, but deviations from the corresponding orthorhombic space group $Ccmm$ ($a = 7.87$, $b = 16.62$, $c = 5.66$ Å) are minimal¹¹. We have used orthorhombic symmetry in our calculations and representations. Using the JEOL ARM-1000 at the National Center for Electron Microscopy, we obtained focus series of images at atomic resolution in all three main directions. We also used the high-angle tilt stage ($\pm 40^\circ$ on two axes) to obtain images perpendicular to [101] and [310]. We selected only micrographs in very thin areas (< 40 Å) close to Scherzer focus for analysis.

An example is shown in Fig. 1*a*. In these thin areas, the scattering contribution from the amorphous film is considerable, but the crystalline component is resolved with a good signal-to-noise ratio. Images in each of the five projections were digitized and Fourier transformed, and the periodic structure reconstructed by inverse Fourier transformation of structure factors determined at the reciprocal lattice points⁶ (Fig. 1*b*). Multibeam dynamical scattering image-contrast calculations¹² for the microscope conditions¹³ and specimen parameters¹¹ agree well with the observed image (Fig. 1*c*). The same is true for diffraction patterns.

Figure 2*a, b* compares the experimental electron diffraction pattern with the calculated pattern. Also shown are a diffractogram obtained by optical diffraction from an experimental image and a diffractogram calculated for a 30-Å-thick specimen (Fig. 2*c, d*). The optical diffractogram shows intensity to 1.4 Å, which is the value at which the envelope of the contrast transfer function of the instrument approaches zero¹³. We used only reflections with $d > 1.6$ Å in the present analysis because of the uncertainty in determining the sign of the contrast transfer function at high resolution. Symmetry-related reflections within each projection were averaged, then equivalent reflections in different projections were averaged, resulting in measurements of 30 of the 80 unique, non-extinct reflections, out to $d = 1.6$ Å.

Projections in five directions were combined to obtain data in the three-dimensional reciprocal space. For the reconstruction of the crystal electron potential $V(\vec{x})$, we applied the equation

$$V(\vec{x}) = \sum_{\vec{g}} |F_{\vec{g}}| e^{i\phi_{\vec{g}}} e^{-i2\pi\vec{g}\cdot\vec{x}}$$

where $|F_{\vec{g}}|$ is the amplitude of the structure factor, $\phi_{\vec{g}}$ the phase, and the summation is over all reciprocal lattice vectors $\vec{g}$. Because all five of the projections of staurolite that we used are centrosymmetric, phases are either 0° or 180° . They are determined directly from the Fourier transform of the image after shifting to the proper phase origin⁶. Amplitudes $|F_{\vec{g}}|$ can be obtained either directly from the electron diffraction pattern (Fig. 2*a*) or from the transform of the image (Fig. 2*c*). We have used the latter. As electron diffraction patterns average over large areas of varying thickness, it is advantageous to obtain amplitude information from images, even though they are affected by the contrast transfer function. The effect of this function on the amplitudes could be corrected in the image transform, but we have not done so here because the main

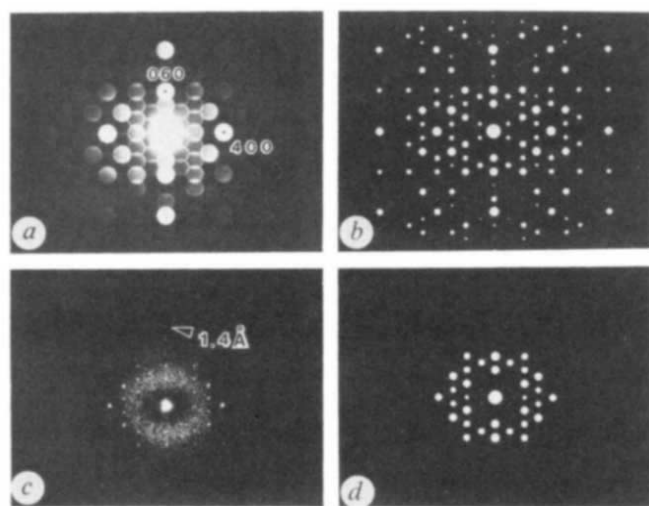


FIG. 2 (*hko*) diffraction patterns of staurolite. *a*, Experimental selected-area diffraction pattern. *b*, Corresponding calculated pattern. *c*, Experimental optical diffractogram from the image in Fig. 1*a* displaying intensity to 1.4 Å (arrow). Rings are due to amorphous contribution. *d*, Calculated diffractogram; same conditions as for Fig. 1*c*.

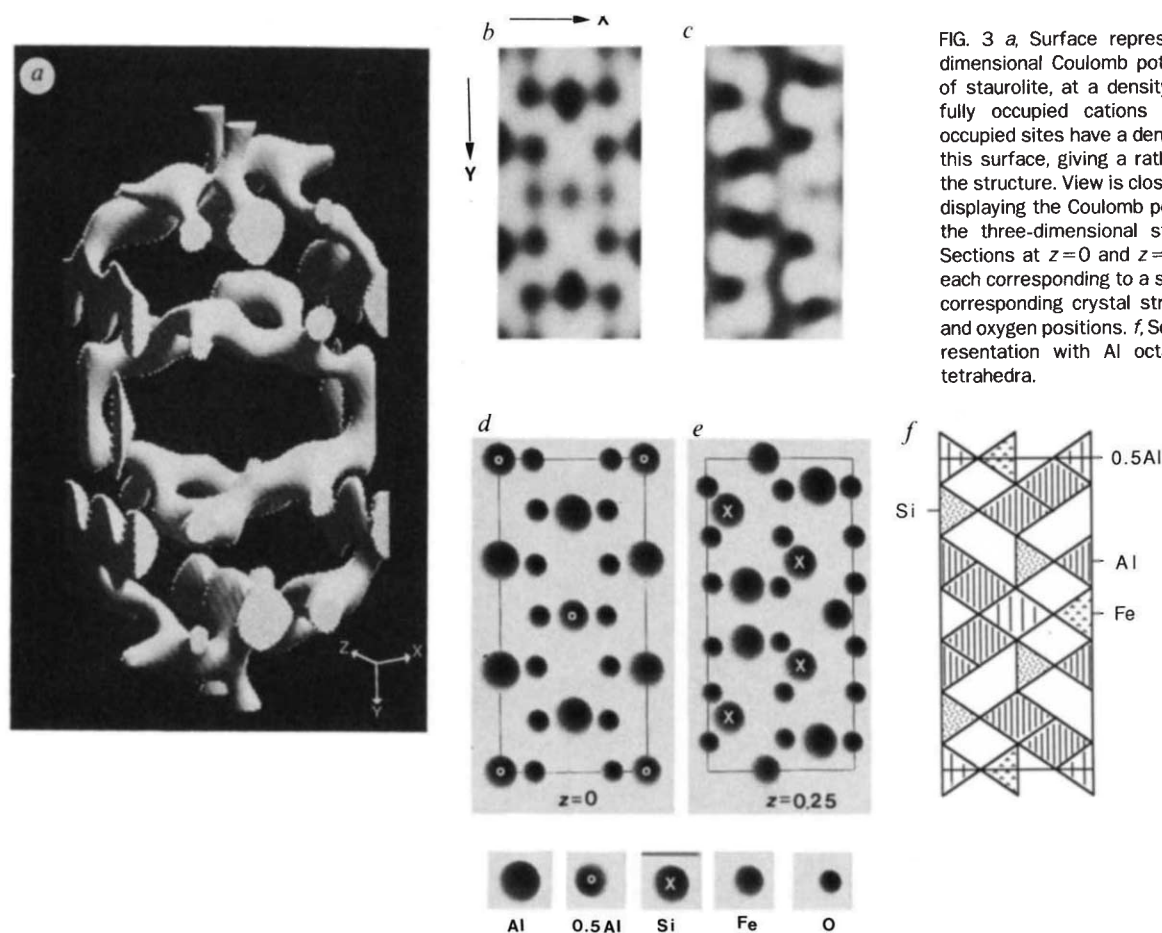


FIG. 3 *a*, Surface representation of the three-dimensional Coulomb potential in a full unit cell of staurolite, at a density level that displays all fully occupied cations and oxygens. Partially occupied sites have a density below the cutoff for this surface, giving a rather open appearance to the structure. View is close to [001]. *b*, *c*, Sections displaying the Coulomb potential as derived from the three-dimensional structure reconstruction. Sections at $z=0$ and $z=0.25$ (units) are shown, each corresponding to a slice 0.2 Å thick. *d*, *e*, The corresponding crystal structure indicating cation and oxygen positions. *f*, Schematic polyhedral representation with Al octahedra and Si and Fe tetrahedra.

structural features rely much more on phases than on amplitudes. Calculations for staurolite show that amplitudes vary linearly with thickness up to 50 Å; above that, changes are irregular owing to dynamic scattering.

In the centrosymmetric case, phases derived from the images are less affected by dynamical effects. Table 1 compares, for some strong reflections of staurolite, calculated amplitudes and phases with observed amplitudes and phases. Measured phases deviate from 0° or 180° as a result of the finite signal-to-noise ratio as well as slight misalignment of the electron beam and bending of the sample which cause deviations from centrosymmetry in the image. All of the measured phases, except for a few weak reflections, were within 30° of 0° or 180°, and when the centrosymmetry constraint was imposed, agreed with phases calculated from the known structure. Deviations between measured and observed amplitudes are larger, but as amplitudes

mainly determine fine details in refinements, the effects of these deviations are too small to cause concern at this stage.

Symmetry operators expanded 30 structure factors to 72 in half-space. With those we calculated a three-dimensional potential map. This is illustrated in Fig. 3*a* as a three-dimensional surface enclosing all of the cations and associated oxygens that are present at full occupancy. Because all atoms are located near $z=0$ and $z=0.25$, and because of the assumed orthorhombic symmetry, most of the information is contained in these two xy sections. Figure 3*b-e* compares sections through the potential map with sections through the structure model. There is excellent correspondence between the experimental potential map and the crystal structure, with all atoms clearly resolved. The partially occupied octahedral Al site at $\frac{1}{2}\frac{1}{2}0$ has a much lower potential than the fully occupied sites ($\frac{1}{2}\frac{1}{2}0$). What surprises us most is the high resolution of light atoms such as oxygen, which can be clearly distinguished. The fact that individual atoms in a close-packed structure can be separated is largely attributed to the three-dimensional reconstruction. We see great potential for three-dimensional electron crystallography in the determination of unknown crystal structures, particularly where homogeneous regions exist only in sub-micrometre-sized domains. Such heterogeneous crystals have been increasingly recognized in metals, ceramics and minerals. We estimate that a three-dimensional structure determination should be possible on areas only about 10 unit cells wide, provided that a sufficient number of projections can be recorded.

TABLE 1 Structure factors for 10 strong reflections of staurolite

<i>h k l</i>	<i>d</i> (Å)	Calculated		Experimental single reflection	
		Amplitude	Phase	Amplitude	Phase
0 6 0	2.77	1.000	0°	1.000	9.1°
0 2 2	2.68	0.277	180°	0.308	169.8°
0 4 2	2.34	0.113	0°	0.186	17.4°
0 6 2	1.98	0.505	0°	0.361	-10.1°
2 0 2	2.30	0.344	180°	0.309	-177.9°
3 1 0	2.59	0.263	180°	0.627	171.7°
3 3 0	2.37	0.618	0°	1.228	-9.9°
3 5 0	2.06	0.060	180°	0.144	159.9°
2 6 0	2.26	0.400	180°	0.460	174.8°
1 3 2	2.40	0.832	0°	1.092	-11.2°

A complete table is available from the authors.

Received 10 July; accepted 17 October 1990.

1. Hashimoto, H., Endoh, H., Tanji, T., Ono, A. & Watanabe, E. *J. phys. Soc. Japan* **42**, 1073-1074 (1977).
2. Smith D. J. *et al. Ultramicroscopy* **18**, 63-76 (1985).
3. Gronsky, R. *Proc. 38th A. Mtg. Electron Microscopy Soc. Am.*, 2-5 (Claitors, Baton Rouge, 1980).
4. Veblen, D. R. & Buseck, P. R. *Science* **206**, 1398-1400 (1979).
5. O'Keefe, M. A., Buseck, P. R. & Iijima, S. *Nature* **274**, 322-324 (1978).
6. Amos, L. A., Henderson, R. & Unwin, P. N. T. *Prog. Biophys. molec. Biol.* **39**, 183-231 (1982).

7. Glaeser, R. M. *A. Rev. phys. Chem.* **36**, 243–275 (1985).
8. Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
9. Henderson, R. *et al. J. molec. Biol.* **213**, 899–929 (1990).
10. Hövemöller, S., Sjögren, A., Farrants, G., Sundberg, M. & Marinder, B. O. *Nature* **311**, 238–241 (1984).
11. Smith, J. V. *Am. Miner.* **53**, 1139–1155 (1968).
12. Kilaas, R. *Proc. 45th A. Mtg Electron Microscopy Soc. Am.* 66–69 (San Francisco Press, 1987).
13. Hetherington, C. J. D., Nelson, E. C., Westmacott, K. H., Gronsky, R. & Thomas, G. *Proc. Symp. 139 Mater. Res. Soc.* (eds Ponce, F. A. & Smith, D. J.) 277–282 (1989).

ACKNOWLEDGEMENTS. We thank the NIH (K.H.D.), the NSF (H.R.W.), the University of California Education Abroad Program (H.M. and H.R.W.) and the U.S. Department of Energy (K.H.D. and M.A.O'K.) for support.

Evidence for multinuclear metal-ion complexes at solid/water interfaces from X-ray absorption spectroscopy

C. J. Chisholm-Brause, P. A. O'Day, G. E. Brown Jr & G. A. Parks

Aqueous and Surface Geochemistry Group, School of Earth Sciences, Stanford University, Stanford, California 94305-2115, USA

METALS dissolved in natural waters often become sorbed onto oxide or clay minerals, so that prediction of their chemical behaviour and transport properties requires knowledge of the structure and bonding of metal species at the solid/water interface¹. For many sorption systems, X-ray absorption spectroscopy (XAS) can be used to determine the identity and number of nearest-neighbour atoms and interatomic distances in aqueous complexes on solid surfaces, and thus to identify the dominant type of surface complex and the partitioning mechanism². Here we describe an XAS study of divalent cobalt (Co(II)) complexes sorbed on three different solids, γ -Al₂O₃, rutile (TiO₂) and kaolinite (Al₂Si₂O₅(OH)₄). We find direct evidence for the presence of multinuclear sorption complexes at surface coverages below one monolayer of Co(II) atoms. Our spectroscopic data reveal distinct differences in the number of coordinating atoms and interatomic distances in the surface complexes formed on each of the solids at the same sorption density. These results suggest that different oxide and clay surfaces influence the structure and properties of aqueous surface complexes, and therefore must be accounted for in models of metal-ion sorption.

Solution experiments^{3–6} and spectroscopic studies^{7–9} have provided indirect evidence for the formation of multinuclear surface complexes or surface precipitates on solid oxides. These types of studies often cannot distinguish unambiguously among different types of surface complexes or sorption mechanisms, such as inner-sphere and outer-sphere complexes, mononuclear and multinuclear complexes, surface precipitates and solid solutions^{9–11}. By identifying the type (either atoms from solution or from the solid) and the number of atoms in the first, second and greater coordination shells around the sorbed atom, its bonding environment on different solids, and thus the sorption mechanism(s), can be characterized. Here we use experimental methods developed specifically for the study of sorption complexes *in situ*^{12–15} with XAS, in particular extended X-ray absorption fine structure (EXAFS) spectroscopy^{16,17}, to determine atom identity, average interatomic distance R , coordination number CN and relative Debye–Waller factor $\Delta\sigma^2$ for Co(II) sorbed on three crystalline solids.

We obtained X-ray absorption spectra at the Stanford Synchrotron Radiation Laboratory and the National Synchrotron Light Source for Co sorbed on γ -Al₂O₃, rutile and kaolinite. Spectra were also obtained for a 12mM Co(NO₃)₂ aqueous solution and for five solid reference ('model') compounds (verified by powder X-ray diffraction) with well-known crystal structures and compositions: two synthetic compounds, Co(OH)₂(s) and [Co(H₂O)₆](ClO₄)₂, and three minerals,

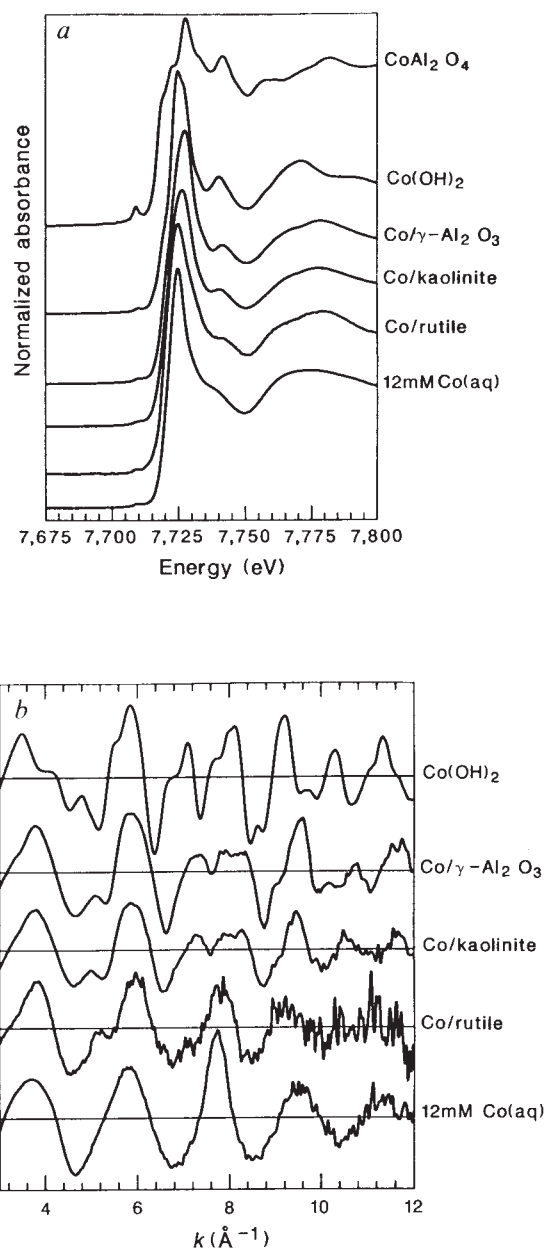


FIG. 1 Normalized background-subtracted Co X-ray absorption spectra for a solid Co(OH)₂ model compound, Co sorbed on γ -Al₂O₃, kaolinite and rutile (1.1–1.2 $\mu\text{mol m}^{-2}$) and for an aqueous 12mM Co(NO₃)₂ solution. *a*, Edge spectra of the above samples compared to that of CoAl₂O₄(s). The pre-edge feature at $\sim 7,709$ eV is small in the spectra of Co(OH)₂, the sorption samples and the aqueous solution, but much larger in the spectrum of CoAl₂O₄ which contains tetrahedrally coordinated Co. *b*, EXAFS spectra for the sorption samples and model compounds, multiplied by k^3 to provide equal weighting of the EXAFS oscillations over the entire k range. Lower signal-to-noise in the rutile sorption sample arises from a high background fluorescence from Ti.

METHODS. γ -Al₂O₃ ('Alon' from Cabot Corp.) is a well characterized synthetic oxide. It is a good analogue for natural aluminium oxides and hydroxides¹⁴. Synthetic rutile was from British Titan Products, and is of high chemical purity and high surface area. The two materials had BET surface areas of 117 and 23 m² g⁻¹, respectively. The kaolinite is a <2 - μm fraction of a natural sample from Dry Branch, Georgia²⁹ and has a surface area of 11 m² g⁻¹. The sorption samples were prepared by equilibrating individual batches of each solid with 1.4–1.7 mM Co(NO₃)₂ solutions (ionic strength 0.1M NaNO₃) for 16–28 h under N₂. The solids concentrations and pH values were selected to achieve $\sim 95\%$ uptake and the same sorption density on each sample (1.1–1.2 $\mu\text{mol m}^{-2}$). Equilibrated samples were centrifuged at 12,000 r.p.m. for 20–60 min and $>90\%$ of the supernatant solution was removed before XAS data collection. The wet pastes were sealed in mylar-
teflon holders under N₂.

TABLE 1 Structural information

	First shell: Co—O			Second shell: Co—Co			Second shell: Co—M _s *			Fourth shell: Co—Co		
	R (Å)	CN	$\Delta\sigma^2$ (Å ²)	R (Å)	CN	$\Delta\sigma^2$ (Å ²)	R (Å)	CN	$\Delta\sigma^2$ (Å ²)	R (Å)	CN	$\Delta\sigma^2$ (Å ²)
Co(OH) ₂	2.10	6.0	—	3.17	6.0	—	—	—	—	6.35	6.0	—
Co- γ -Al ₂ O ₃	2.08	6.0	0.0004	3.14	3.0	0.0007	3.27	1.7	0.0026	6.08	1.2	-0.0056
Co-kaolin	2.08	6.0	0.0032	3.14	2.2	-0.0011	3.25	0.6	0.0018	—	—	—
Co-TiO ₂	2.09	6.0	0.0020	3.33	0.8	0.0010	3.29	1.6	0.0031	—	—	—
Co ²⁺ (aq)	2.10	6.6	-0.0025	—	—	—	—	—	—	—	—	—

R, CN and $\Delta\sigma^2$ for the sorption samples and aqueous solution were determined by least-squares fits of the Fourier back-transformed EXAFS functions of selected peaks of the radial structure functions in Fig. 2. The atomic composition of each shell was varied during the analysis to determine which atoms resulted in the best fit of the data. For forward Fourier transforms, limits on *k* ranged from ~ 3 to $12 \text{ \AA}^{-1}$. For back-transforms, *R* ranges varied slightly depending on peak position: first shell, 0.5–0.6 to 2.0–2.1 Å; second shell, 2.1–2.3 to 3.1–3.5 Å; fourth shell, 5.1–5.3 to 6.1–6.4 Å. The distances reported above differ from those observed in the radial structure functions because the latter have not been corrected for phase shift. *R* and CN for Co(OH)₂ are from X-ray diffraction measurements²⁸. $\Delta\sigma^2$ is the difference between the σ^2 parameter derived from fits of model compounds and σ^2 calculated from fits of the unknown samples.

*M_s represents Al for Co sorbed to γ -Al₂O₃, Al or Si for Co sorbed on kaolinite, and Ti for Co sorbed on rutile.

ilmenite (FeTiO₃), andradite (Ca₃Fe₂(SiO₄)₃) and hedenbergite (CaFeSi₂O₆).

We collected the fluorescence-yield XAS data at room temperature under He for wet sorption samples and for the aqueous solution using standard methods^{18,19}; spectra for model compounds were collected in transmission mode. Multiple scans (4–10) for each sorption sample were averaged to increase the signal-to-noise ratio. Spectra were calibrated by assigning the first inflection point in the K-edge spectrum of Co foil (collected simultaneously) to 7,709.25 eV. Data reduction consisted of pre-edge and background subtraction and normalization of the edge jump using McMaster coefficients to yield a normalized spectrum²⁰ (Fig. 1a). After conversion of the energy scale to a wavevector scale ($k = \{8\pi^2 m_e (E - E_0)/h^2\}^{1/2}$, where m_e is the mass of the electron, *E* is the energy, *E*₀ the energy at *k* = 0, and *h* is Planck's constant), the resultant EXAFS spectra (Fig. 1b) were Fourier transformed to yield radial structure functions consisting of *n* peaks representing *n* coordination shells at distances *R*_{*n*}, uncorrected for phase shift^{17,21} (Fig. 2). We identified spurious peaks in the radial structure functions owing to noise and Fourier termination effects by noting variations in their amplitude and position with changes in the *k* limits of the transform. Peaks resulting from first-, second- and in some spectra fourth-shell Co/backscatterer atom interactions were separately back-transformed to yield filtered EXAFS functions, each representing only the phase and amplitude contributions from atoms in that coordination shell.

We derived phase shift and amplitude functions characteristic of Co—O, Co—Co, Co—Ti and Co—Al or Co—Si pairs in bonding environments similar to those expected in the sorption samples by empirical fits of filtered EXAFS spectra for the solid model compounds. The accuracy of the fit parameters was estimated by treating one model as an unknown; for example, calculating *R*, CN, $\Delta\sigma^2$ and *E*₀ for Co—O in [Co(H₂O)₆](ClO₄)₂ using parameters derived from Co(OH)₂(s) and *vice versa*. This procedure and other studies of empirical fitting of model compounds^{15,22} indicate errors of $\pm 0.01 \text{ \AA}$ in *R* and of $\pm 10\%$ in CN for first-shell atoms, and of $\pm 0.02 \text{ \AA}$ and $\pm 25\%$ for second-shell atoms. Phase and amplitude functions for elements of atomic number *Z* differing by ± 1 are not significantly different¹⁷. Therefore, when slight differences in *E*₀ are accounted for in the fitting procedure, Fe—Ti functions (from ilmenite) can be used for Co—Ti and Fe—Si functions (from andradite and hedenbergite) can be used for both Co—Al and Co—Si functions. Unfortunately, this similarity makes Si and Al indistinguishable in EXAFS analysis. The differences in *Z* between Co and atoms from the solids (M_s = Al, Si or Ti), however, are large enough that Co—Co and Co—M_s pairs can be distinguished, allowing identification of Co, Ti or Si or Al in the second coordination shell.

The lack of a significant pre-edge feature at 7,709 eV (arising from a 1s to 3d electronic transition) in the normalized edge spectra (Fig. 1a) suggests that the Co site is nearly centrosym-

metric¹⁸. Comparison of normalized EXAFS spectra (Fig. 1b) shows distinct differences in the frequency and amplitude of oscillations between the sorption samples and either Co²⁺(aq) or solid Co(OH)₂. A 'beat pattern' (seen as lower-amplitude shoulders at *k* = 4.8–5.2, 7.0–7.4 and 8.9–9.2 Å⁻¹ in the sorption spectra) due to backscattering from second-neighbour atoms has similar amplitude in the spectra of Co/ γ -Al₂O₃ and Co/kaolinite, a lower amplitude in the Co/rutile spectrum, and is absent in the Co²⁺(aq) spectrum. Nonlinear least-squares fits of filtered EXAFS spectra using empirical phase shift and amplitude functions derived from the model compounds²³ yielded first-, second- and fourth-shell atom identities, CN, *R* and $\Delta\sigma^2$ (Table 1), confirming the presence and types of neighbouring atoms suggested by the normalized EXAFS spectra.

The sorption complexes on all three solids have similar first-coordination shells consisting of six oxygens at an average

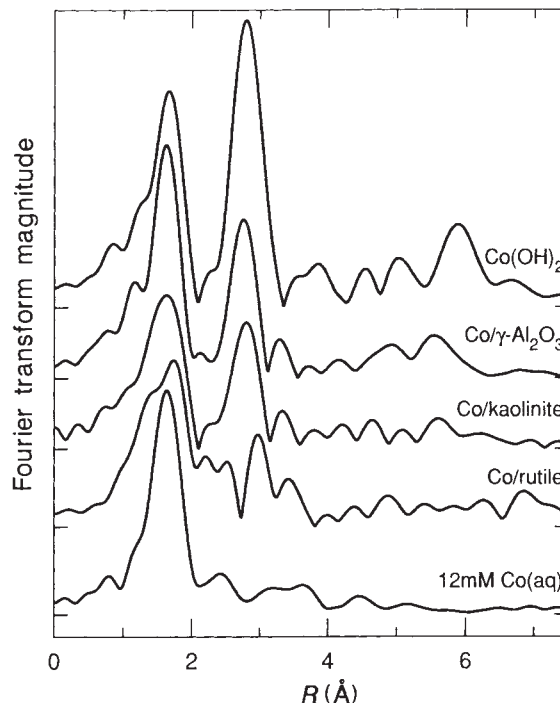


FIG. 2 Radial structure functions, uncorrected for phase shift, for Co(OH)₂(s), Co sorbed on γ -Al₂O₃, kaolinite and rutile (1.1–1.2 $\mu\text{mol m}^{-2}$) and aqueous 12mM Co(NO₃)₂ solution. The first major peak is attributed to first-neighbour oxygen atoms and its amplitude reflects the number of first-shell oxygens surrounding Co. The second peak for Co(OH)₂ and the sorption samples is attributed to second-shell atoms surrounding the Co. The amplitude of this peak is lower in the sorption samples relative to Co(OH)₂, and it is absent in the aqueous cobalt solution. A peak at $\sim 6 \text{ \AA}$ owing to more distant Co atoms is present for Co(OH)₂ and, with lower amplitude, for Co sorbed on γ -Al₂O₃, but is not above background for the other samples.

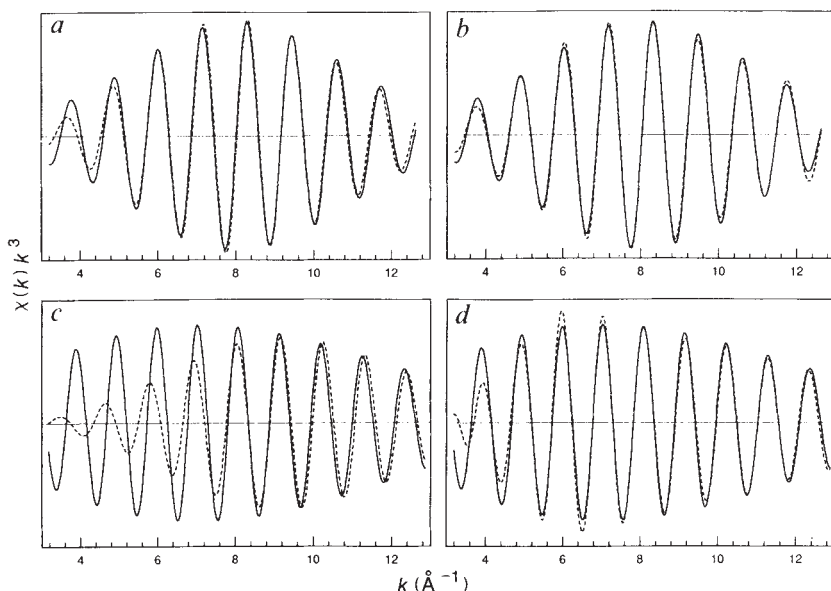


FIG. 3 EXAFS function generated by Fourier back-transforming the second peak of the radial structure function for Co sorbed on γ - Al_2O_3 or rutile (solid lines) compared with least-squares fits (dashed lines) with the following second-neighbour atoms: a, only Co in the second coordination shell of the central Co sorbed on γ - Al_2O_3 ; b, Co and Al in the second shell of Co sorbed on γ - Al_2O_3 ; c, only Co in the second shell of Co sorbed on rutile; d, Co and Ti in the second shell of Co sorbed on rutile. For all sorption-sample spectra, significantly better fits were obtained when both Co and metal atoms of the oxide surface (Ti, Al or Si) were included in the second shell. Fits for kaolinite are nearly identical to those for γ - Al_2O_3 and are not shown.

distance of 2.08–2.09 Å (Table 1). This CN and the low intensity of the pre-edge feature imply octahedral coordination of Co. The amplitude of the second peak in the radial structure function (Fig. 2) decreases in the order $\text{Co}(\text{OH})_2(\text{s}) > \text{Co}/\gamma\text{-Al}_2\text{O}_3 > \text{Co}/\text{kaolinite} > \text{Co}/\text{rutile}$; $\text{Co}^{2+}(\text{aq})$ has no second peak. Best fits of the second shell yield Co coordination numbers that also decrease in this order (Table 1). Co–Co distances for sorption complexes on γ - Al_2O_3 and kaolinite ($R = 3.14$ Å) are similar to those in $\text{Co}(\text{OH})_2(\text{s})$ ($R = 3.17$ Å), but these distances for complexes on rutile are 0.19 Å longer (Table 1). Fits of the second peak for the sorption samples are significantly improved by including atoms from the solid ($M_s = \text{Al}$, Si or Ti) as well as Co in the second shell around Co (Fig. 3). The small number of M_s atoms (0.6–1.7) suggests that Co diffusion into the solid is not significant. The presence of second-neighbour Co indicates a significant fraction of the sorbed Co forms multinuclear sorption complexes and/or a surface precipitate at the solid/water interface. The small number of second-shell Co obtained for the sorption samples relative to $\text{Co}(\text{OH})_2(\text{s})$ (Table 1), however, indicates that most of the Co forms multinuclear complexes. Furthermore, the radial structure function for Co sorbed on γ - Al_2O_3 has a small peak at ~ 5.7 Å (Fig. 2) which is caused by multiple scattering among equidistant collinear Co atoms²⁴ in the fourth coordination shell about a central Co, as inferred from a similar feature at 5.9 Å for $\text{Co}(\text{OH})_2(\text{s})$. This peak is not above background in the radial structure function for Co sorbed to rutile or kaolinite at this surface coverage, suggesting longer-range Co interactions and larger multinuclear sorption complexes on γ - Al_2O_3 than on rutile or kaolinite. This peak grows in amplitude and CN increases at higher surface coverages and in samples where $\text{Co}(\text{OH})_2(\text{s})$ solubility was exceeded²⁵; the radial structure functions of these samples are not, however, identical to that of $\text{Co}(\text{OH})_2(\text{s})$, which suggests the lack of a well ordered three-dimensional precipitate of $\text{Co}(\text{OH})_2(\text{s})$. The small values obtained for $\Delta\sigma^2$ (Table 1) suggest a low degree of static and thermal disorder in the local environment of sorbed and aqueous Co relative to the solid model compounds.

Our data indicate that a majority of the Co sorbed on γ - Al_2O_3 , kaolinite and rutile under our experimental conditions forms multinuclear complexes in which each Co is octahedrally coordinated by oxygens and one to three second-neighbour Co atoms. The presence of second-shell metal atoms from the solid at 3.25–3.29 Å indicates that Co loses part of its hydration sphere and bonds directly to oxygens of the solid surface forming an inner-sphere complex, in agreement with results from previous

solution studies²⁶. For similar solution conditions and surface coverages, however, the solid affects the number of atoms in, and the structure of, the surface complex as shown by the progressive decrease in the number of second-neighbour Co atoms on the three solids, the absence of more distant Co neighbours on kaolinite and rutile and their presence on γ - Al_2O_3 , and the increase in second-shell interatomic distances for Co on rutile relative to γ - Al_2O_3 and kaolinite.

These results provide direct evidence for multinuclear metal sorption complexes at solid/water interfaces, confirming hypotheses based on earlier, less-direct evidence in a variety of other aqueous sorption systems. The presence of multinuclear sorption complexes is not wholly surprising because multinuclear Co hydroxo complexes form in solution²⁷. But such aqueous species are present only in negligible concentrations in our system²⁷, and they are therefore unlikely to be the predominant species adsorbing directly from solution. These complexes are also unlikely to be the result of a random two-dimensional distribution of Co atoms at the sorption density of our experiments (1.1 – $1.2 \mu\text{mol m}^{-2}$), which would result in an average Co–Co distance (11.2 Å) much longer than that determined in this study (3.14 – 3.33 Å). Multinuclear sorption complexes are probably more common than has been expected. Their presence and the dependence of their structures and compositions on the nature of the adsorbent must be explicitly accounted for in realistic thermodynamic and kinetic descriptions of partitioning processes. □

Received 25 June; accepted 18 October 1990.

1. Parks, G. A. *Chem. Aust.* **49**, 389–395 (1981).
2. Sposito, G. *Chimia* **43**, 169–176 (1989).
3. James, R. O. & Healy, T. W. *J. Colloid Interface Sci.* **40**, 53–64 (1972).
4. Matijevic, E. *J. Colloid Interface Sci.* **43**, 217–245 (1983).
5. Tripathi, V. S. thesis, Stanford Univ. (1983).
6. Farley, K. J., Dzombak, D. A. & Morel, F. M. M. *J. Colloid Interface Sci.* **106**, 226–242 (1985).
7. Tewari, P. H. & Lee, W. J. *Colloid Interface Sci.* **52**, 77–88 (1975).
8. Bleam, D. F. & McBride, M. B. *J. Colloid Interface Sci.* **103**, 124–132 (1985).
9. Motoshi, H. in *Aquatic Surface Chemistry* (ed. Stumm, W.) 111–125 (Wiley, New York, 1987).
10. Sposito, G. *Surface Chemistry of Soils* (Oxford University Press, New York, 1984).
11. Davis, J. A., James, R. O. & Leckie, J. O. *J. Colloid Interface Sci.* **63**, 480–499 (1978).
12. Brown, G. E. Jr & Parks, G. A. *Rev. Geophys.* **27**, 519–533 (1989).
13. Brown, G. E. Jr, Parks, G. A. & Chisholm-Brause, C. J. *Chimia* **43**, 248–256 (1989).
14. Chisholm-Brause, C. J. et al. *Geochim. cosmochim. Acta* **54**, 1897–1909 (1990).
15. Hayes, K. F. et al. *Science* **238**, 783–786 (1987).
16. Teo, B.-K. & Lee, P. A. *J. Am. chem. Soc.* **101**, 2815–2832 (1979).
17. Stern, E. A. in *X-ray Absorption* (eds Kroningsberger, D. C. & Prins, R.) 3–51 (Wiley, New York, 1988).
18. Brown, G. E. Jr, Calas, G., Waychunas, G. A. & Petiau, J. in *Spectroscopic Methods in Mineralogy and Geology* (ed. Hawthorne, F. C.) 431–512 (Miner. Soc. Am., Chelsea, Michigan, 1988).
19. Heald, S. M. in *X-ray Absorption* (eds Kroningsberger, D. C. & Prins, R.) 87–118 (Wiley, New York, 1988).
20. Cramer, S. P. & Hodgson, K. O. *Prog. inorg. Chem.* **25**, 1–39 (1979).

21. Sayers, D. C. & Bunker, B. A. in *X-ray Absorption* (eds Kroningsberger, D. C. & Prins, R.) 211–253 (Wiley, New York, 1988).
22. Waychunas, G. A., Brown, G. E. Jr & Apter, M. J. *Phys. Chem. Miner.* **13**, 31–47 (1986).
23. Cramer, S. P., Hodgson, K. O., Stiefel, E. I. & Newton, W. E. *J. Am. chem. Soc.* **100**, 1144–1150 (1983).
24. Co, M. S., Hendrickson, W. A., Hodgson, K. O. & Doniach, S. *J. Am. chem. Soc.* **105**, 2748–2761 (1978).
25. O'Day, P. A., Brown, G. E. Jr & Parks, G. A. in *Proc. 6th Int. Conf. on X-ray Absorption Fine Structure* (ed. S. S. Hasnain) (Ellis-Horwood Publishers, in the press).
26. Hachiya, K., Sasaki, M., Saruta, Y., Mikami, N. & Yasunaga, T. *J. phys. Chem.* **88**, 23–27 (1984).
27. Baes, C. F. Jr & Mesmer, R. E. *The Hydrolysis of Cations* (Wiley, New York, 1986).
28. Lotmar, V. W. & Feitknecht, W. Z. *Kristallogr.* **A93**, 368–378 (1936).
29. May, H. M., Kinniburgh, D. G., Helmke, P. A. & Jackson, M. L. *Geochim. cosmochim. Acta* **50**, 1667–1677 (1986).

ACKNOWLEDGEMENTS. We thank R. O. James (Eastman Kodak) and H. M. May (US Geological Survey) for well characterized samples of rutile and kaolinite, respectively. We also thank A. L. Roe (Intel Corp.) for help with data collection on the model compounds and K. O. Hodgson (Stanford) for use of his computer facilities. We acknowledge the NSF for financial support. The Stanford Synchrotron Radiation Laboratory is supported by DOE and NIH.

Role of fluids in transport and fractionation of uranium and thorium in magmatic processes

Hans Keppler* & Peter J. Wyllie

California Institute of Technology, Division of Geological and Planetary Sciences, 170–25, Pasadena, California 91125, USA

THE geochemistry of uranium and thorium is of considerable importance for understanding the Earth's heat budget^{1,2}, for U–Th–Pb age determinations³, for the origin of ore deposits^{4–6}, and because disequilibria between the radioactive daughters of ²³⁸U and ²³²Th provide constraints on processes occurring during the generation of magmas^{7–14}. We have studied experimentally the partitioning of uranium and thorium between a haplogranitic melt and aqueous fluids containing variable amounts of HCl, HF and CO₂, at 2 kbar, 750 °C and the oxygen fugacity of the Ni–NiO buffer. The partition coefficients $K_D^{\text{fluid/melt}}$ are very low for both uranium and thorium if water is the only volatile component present, but they increase strongly with increasing fluoride concentration, indicating the formation of fluoride complexes in the fluid. Chloride and CO₂, on the other hand, form complexes with uranium, but not with thorium. These results explain the origin of hydrothermal uranium and thorium deposits, the fractionation of uranium from thorium during magma formation, and the depletion of uranium relative to thorium in granulite-facies rocks.

Although there are several experimental studies on the partitioning of chalcophile trace elements such as Cu, Mo, Pb and Zn^{15,16} in magmatic-hydrothermal systems, virtually no experimental evidence on the behaviour of U and Th is available. Webster *et al.*¹⁷ reported some data on U and Th partitioning, but their results must be treated with caution, because of the presence of zircon in the glasses analysed.

We studied the partitioning of U and Th between melt and aqueous fluids in the system haplogranite–H₂O–HCl, haplogranite–H₂O–HF and haplogranite–H₂O–CO₂ at 2 kbar and 750 °C using rapid-quench cold-seal autoclaves. Starting materials were a dehydrated gel (35 wt% SiO₂, 40 wt% NaAlSi₃O₈, 25 wt% KAlSi₃O₈, doped with 0.1 or 0.5 wt% Th as ThO₂ or with 0.5 wt% U as U₃O₈) and aqueous solutions of HCl and HF. Ag₂C₂O₄ was used for generation of CO₂. The solid and solution (60 mg of each) were loaded into a Pt capsule (25 mm long, 3.8 mm in diameter, with walls 0.1 mm thick). This capsule was welded shut and sealed into a larger Au capsule (30 mm long, 5 mm in diameter, with walls 0.1 mm thick) together with Ni powder, NiO and H₂O. Run durations were 2–6 weeks. After rapid quenching, the solutions were withdrawn with a microsyringe and the sample boiled in concentrated HCl to redissolve products precipitated from the aqueous fluid during

quenching. This extract was added to the withdrawn solution. Both the quenched melt and the fluid phase were analysed by plasma emission spectrometry (DCP).

In experiments with H₂O as the only volatile, the U content in the solution measured after acid treatment of the run products was below detection limit. This shows that no significant amount of U was leached out of the glass. This probably also holds for Th, as Th compounds are generally less soluble than compounds of U. The glasses were usually free of crystals, only traces of quartz were occasionally observed. Bubbles present in the glass might have formed during quenching, and mass balance calculations show that the amount of fluid trapped in them is too small to significantly affect the measurement.

Several sets of reversed experiments were performed. In one set of these runs, U- and Th-doped glasses (containing 0.1 wt% ThO₂ or 0.5 wt% U₃O₈ dissolved in the glass matrix) and solutions free of U and Th were used as starting materials. In the other set of experiments, all U and Th was loaded as a solution containing 0.1 wt% Th as Th(NO₃)₄ or 0.5 wt% U as UO₂(NO₃)₂. The two different starting materials always yielded very similar partition coefficients, demonstrating attainment of equilibrium.

Experimental results for the partitioning of U in the systems haplogranite–H₂O–HCl and haplogranite–H₂O–HF are given in Fig. 1a and Table 1. The horizontal axis in this and the following figures gives the composition (in molality) of the solution loaded into the capsule. Although almost all chlorine remains in the aqueous fluid, a considerable fraction of fluorine enters the melt phase. According to our measurements, the partition coefficient K_D of fluorine between melt and aqueous fluid in this system is approximately one. The partition coefficient of U in the presence of H₂O as the only volatile is very low, <0.002, but increases strongly both with increasing HF and increasing HCl

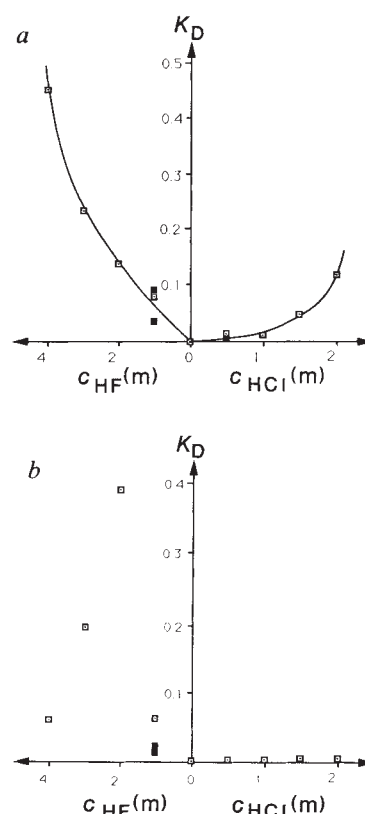


FIG. 1 Partition coefficient $K_D = c_{\text{fluid}}/c_{\text{melt}}$ (in weight units) of a, U and b, Th in the systems haplogranite–H₂O–HCl and haplogranite–H₂O–HF at 2 kbar, 750 °C, Ni–NiO buffer conditions. The horizontal axis gives the molality of the solution loaded into the sample capsule. 0 corresponds to pure water. Filled squares refer to reversed experiments. The results of the reversal at 0.5 m HCl for U are almost identical.

* Present address: Bayerisches Geoinstitut, Universität Bayreuth, W-8580 Bayreuth, Germany.

TABLE 1 Partitioning of U and Th between fluid and haplogranitic melt

Run	Fluid*	Duration (d)	U in fluid (p.p.m.)	U in glass (p.p.m.)	$K_D^{\text{fluid/melt}}$ (U)	Th in fluid (p.p.m.)	Th in glass (p.p.m.)	$K_D^{\text{fluid/melt}}$ (Th)
V65	4 m HF	28	1,100	2,440	0.45			
V64	3 m HF	28	670	2,870	0.23			
V34	2 m HF	24	320	2,310	0.14			
V35	1 m HF	24	230	2,960	0.078			
V96†	1 m HF	41	120	3,190	0.037			
V97‡	1 m HF	41	240	2,580	0.093			
V29	H ₂ O	42	<5	2,860	<0.002			
V62	0.5 m HCl	20	52	3,530	0.015			
V79†	0.5 m HCl	22	17	3,400	0.005			
V81‡	0.5 m HCl	26	15	3,700	0.004			
V30	1 m HCl	42	39	3,010	0.013			
V63	1.5 m HCl	28	140	2,800	0.050			
V31	2 m HCl	36	290	2,380	0.12			
V92	2.5 m CO ₂	19	42	1,910	0.022	6.1	1,020	0.006
V93	5 m CO ₂	19	62	1,880	0.033	2.8	880	0.003
V85	4 m HF	26				68	1,130	0.060
V84	3 m HF	26				100	520	0.19
V71	2 m HF	20				920	2,330	0.39
V83	1 m HF	14				50	800	0.062
V88†	1 m HF	24				13	930	0.014
V89‡	1 m HF	23				21	880	0.024
V74	H ₂ O	14				2.0	740	0.003
V76	0.5 m HCl	18				11	2,550	0.004
V78	1 m HCl	14				2.9	920	0.003
V86	1.5 m HCl	26				5.7	1,120	0.005
V77	2 m HCl	18				17	3,030	0.006

All runs at 2 kbar, 750 °C and Ni–NiO buffer conditions. Concentrations in p.p.m. were calculated as mg kg⁻¹.

* Compositions refer to the solution loaded into the sample capsule at the beginning of the experiment (m = mol kg⁻¹).

† Reversed run, all U or Th was loaded as a solution containing Th(NO₃)₄ or UO₂(NO₃)₂.

‡ Reversed run, all U or Th was loaded as a glass containing dissolved ThO₂ or U₃O₈.

concentration. This observation suggests complexing of U by chlorine and fluorine in the fluid phase. A quantitative analysis of the data shows that K_D of U increases with the square of the fluorine concentration, but with the third power of the chlorine concentration. This is consistent with stoichiometric ratios of F:U = 2:1 and Cl:U = 3:1 in the complexes in the aqueous fluid if only one species is present.

The partition coefficient of Th is ~0.004 in the presence of water as the only volatile component and in the haplogranite–H₂O–HCl system (Fig. 1b, Table 1). No dependence of K_D on HCl concentration is observed. There is, however, clearly an increase in the K_D of Th in the presence of fluorine, although the data show considerable scatter. This scatter is not the result of incomplete equilibration during the run, because reversed experiments give similar results. The scatter could be caused by precipitation of some Th from the fluid during quenching, which could not entirely be redissolved by treatment with HCl. This is possible because of the relatively low solubility of Th fluorides. In this case, the data given in Fig. 1b for the system with fluorine provide only a lower limit for the partition coefficient.

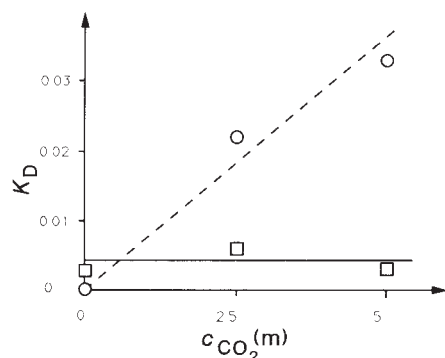


FIG. 2 Partition coefficients $K_D = c_{\text{fluid}}/c_{\text{melt}}$ of U (○) and Th (□) in the system haplogranite–H₂O–CO₂ at 2 kbar, 750 °C, Ni–NiO buffer conditions.

Some preliminary results on the partitioning of U and Th in the system haplogranite–H₂O–CO₂ are given in Fig. 2 and Table 1. Although the K_D of Th is not affected by the presence of CO₂, U seems to form carbonate complexes resulting in an increase of K_D with increasing CO₂ concentration. The high stability of carbonate complexes of U at low temperatures is well established¹⁸, but our data is evidence for their importance in magmatic–hydrothermal processes.

It has been suggested^{4–6} that some U and Th deposits were formed by the partitioning of U and Th into a fluid during the crystallization of a granitic magma. Our experimental results show that such a mechanism is possible, provided that sufficient amounts of fluorine or chlorine are available. Although chloride-containing fluids are capable of transporting only U, both U and Th can be mobilized in the presence of fluoride. This is consistent with the prominent fluorine mineralization associated with hydrothermal deposits containing both U and Th, such as Crockers Well, South Australia⁴. Magmatic–hydrothermal fluids might also have had a role in the formation of the economically important U deposit of Rössing, Namibia⁶. Hydrothermal alteration associated with the emplacement of granitic intrusions can mobilize Pb, U and Th, and therefore disturb the ²³⁸U–²⁰⁶Pb, ²³⁵U–²⁰⁷Pb and ²³²Th–²⁰⁸Pb isotope systems¹⁹. According to our experiments, fluids containing different amounts of fluoride, chloride and CO₂ will have very different effects. Therefore, it should be possible to distinguish the alteration effects of different fluids based on U–Th–Pb isotope systematics. As the composition of fluids controls the enrichment of metals in hydrothermal ore deposits, this could be a useful tool in exploration geology.

High-grade granulite-facies rocks are usually depleted in K, Th and particularly strongly in U (ref. 20). To decide whether this depletion is caused by the extraction of melts or by interaction with fluids, one has to compare the partition coefficients $K_D^{\text{fluid/solid}}$ and $K_D^{\text{melt/solid}}$ of U and Th. The partition coefficients $K_D^{\text{fluid/melt}}$ that we have measured are related to these coefficients by $K_D^{\text{fluid/melt}} = K_D^{\text{fluid/solid}}/K_D^{\text{melt/solid}}$, provided that fluid, melt and solid are in mutual equilibrium. As the haplogranitic melt

composition studied here is similar to that of magmas obtained by partial melting of crustal rocks, the stated relationship holds, at least approximately. Because $K_D^{\text{fluid/melt}}$ is very low for both U and Th in the presence of water as the only volatile (<0.002 and 0.004 , respectively), this means that $K_D^{\text{melt/solid}} > 500 K_D^{\text{fluid/solid}}$ for U, and $K_D^{\text{melt/solid}} = 250 K_D^{\text{fluid/solid}}$ for Th. Melts are therefore much more effective than a pure aqueous fluid in extracting U and Th out of the lower crust. This conclusion changes if complexing ions are present in the fluid. As U but not Th is complexed in the presence of CO_2 , interaction with a CO_2 -containing fluid could account for the depletion of U relative to Th in high-grade metamorphic rocks. The different complexing behaviour of U and Th in fluids also provides an

elegant explanation for the strong fractionation of U and Th during the formation of volatile-rich magmas, as demonstrated by studies of ^{230}Th – ^{238}U disequilibria in young volcanic rocks^{7–14}. These studies generally suggest an enrichment of U over Th in the melt, if fluids are present during magma formation¹⁰. This interpretation is in agreement with our observation that U, but not Th, is complexed by chloride and CO_2 in fluids. If there is an influx of aqueous fluid containing chloride or CO_2 in the source region of a magma, the melt will be enriched in U that was transported by the fluid. In particular, the extraordinarily high enrichment of U over Th observed in carbonatite lavas from Oldoinyo Lengai volcano⁷ emphasizes the importance of carbonate complexes in the transport of U. $\square$

Received 8 August; accepted 11 October 1990.

- Jochum, K. P., Hofmann, A. W., Ito, E., Seufert, H. M. & White, W. M. *Nature* **306**, 431–436 (1983).
- Galer, S. J. G. & O'Nions, R. K. *Nature* **316**, 778–782 (1985).
- Whitehouse, M. J. *Geochim. cosmochim. Acta* **53**, 717–724 (1989).
- Ashley, P. M. *Mineral. Deposita* **19**, 7–18 (1984).
- Simpson, P. R., Brown, G. C., Plant, J. & Ostle, D. *Phil. Trans. R. Soc. A* **291**, 385–412 (1979).
- von Backström, J. W. & Jacob, R. E. *Phil. Trans. R. Soc. A* **291**, 307–319 (1979).
- Williams, R. W., Gill, J. B. & Bruland, K. W. *Geochim. cosmochim. Acta* **50**, 1249–1259 (1986).
- Newman, S., Finkel, R. C. & MacDougall, J. D. *Geochim. cosmochim. Acta* **48**, 315–324 (1984).
- Krishnaswami, S., Turekian, K. K. & Bennett, J. T. *Geochim. cosmochim. Acta* **48**, 505–511 (1984).
- Condomines, M., Hemond, C. & Allegre, C. J. *Earth planet. Sci. Lett.* **90**, 243–262 (1988).
- Villemant, B. & Flehoc, C. *Earth planet. Sci. Lett.* **91**, 312–326 (1989).
- Condomines, M. *et al. Nature* **325**, 607–609 (1987).

- Newman, S., MacDougall, J. D. & Finkel, R. C. *Contrib. Miner. Petrol.* **93**, 195–206 (1986).
- Sigmarsson, O., Condomines, M., Morris, J. D. & Harmon, R. S. *Nature* **346**, 163–165 (1990).
- Candela, P. A. & Holland, H. D. *Geochim. cosmochim. Acta* **48**, 373–380 (1984).
- Urabe, T. *Econ. Geol.* **80**, 148–157 (1985).
- Webster, J. D., Holloway, J. R. & Hervig, R. L. *Econ. Geol.* **84**, 116–134 (1989).
- Langmuir, D. *Geochim. cosmochim. Acta* **42**, 547–569 (1978).
- Silver, L. T. *Epstein 70th Birthday Symp. abstracts with program*, 107–109 (California Institute of Technology, Pasadena, 1989).
- Scharbert, H. G., Korkisch, J. & Steffan, I. *Tschermaks Mineralog. Petrog. Mitt.* **23**, 223–232 (1976).

ACKNOWLEDGEMENTS. We thank E. Stolper for allowing us access to the DCP laboratory and S. Newman for introducing us to its operation. Financial support for this study was provided by the NSF (P.W.) and the NATO/German Academic Exchange Service (H.K.).

The strength of mantle silicates at high pressures and room temperature: implications for the viscosity of the mantle

Charles Meade* & Raymond Jeanloz

Department of Geology and Geophysics, University of California, Berkeley, California 94720, USA

THE rheological properties of the upper mantle, important in convective and tectonic processes, have long been studied through low-pressure creep experiments on olivine^{1–5}. In comparison, our understanding of deformation in the deep Earth is incomplete, because there are few experimental constraints on the rheology of the transition zone (400–670 km depth) and the lower mantle. Estimates of the viscosity at these depths have had to rely on geophysical models of, for example, post-glacial uplift, the geoid and changes in the Earth's rate of rotation^{6–13}. Recently we developed techniques for measuring solid-state creep at mantle pressures¹⁴, and here we use these methods to make direct laboratory measurements of the strength of key mantle minerals. We study $(\text{Mg}, \text{Fe})_2\text{SiO}_4$ olivine (representative of about 60% of the upper mantle), $(\text{Mg}, \text{Fe})_2\text{SiO}_4$ γ -spinel (which constitutes a large fraction of the transition zone) and $(\text{Mg}, \text{Fe})\text{SiO}_3$ perovskite and perovskite + (Mg, Fe) magnesiowüstite assemblages (which comprise most of the lower mantle), all at room temperature and pressures of up to 60 GPa. Our results suggest that the transition zone may form a layer of relatively high viscosity or strength between the upper and lower mantle. A simple rheological model shows that this behaviour may be compatible with current geophysical observations.

For these experiments we used powdered San Carlos olivine ($\text{Mg}_{0.88}\text{Fe}_{0.12}\text{SiO}_4$ (ref. 15) and Bamble enstatite ($\text{Mg}_{0.88}\text{Fe}_{0.12}\text{SiO}_3$ (ref. 16; $\sim 1\text{--}3\text{ }\mu\text{m}$ grain size). These materials were compressed with $<3\%$ weight ruby, of similar grain size, in a Mao–Bell diamond cell (350- μm culets) using spring steel gaskets¹⁴. The pressures were measured with the ruby fluores-

cence technique¹⁷; the ruby was placed at the diamond/sample interface to minimize reactions between the pressure calibrant and the starting materials at high temperatures and to minimize the influence of the ruby on the strength of the sample. For the experiments on perovskite and spinel, the appropriate starting materials (enstatite and olivine, respectively) were compressed well into the stability field of the high-pressure phases at room temperature ($P > 35\text{ GPa}$ for perovskite, $P \approx 20\text{ GPa}$ for spinel). The samples were then heated with Nd:YAG laser radiation to subsolidus temperatures between 1,500–2,000 K. The transformation to the high-pressure phases were confirmed by the significant change in colour (translucent to brown/opaque for perovskite and perovskite + magnesiowüstite; translucent to blue for spinel). X-ray diffraction measurements on each decompressed sample confirmed the phase transformations and provided estimates of the amount of residual starting material in the high-pressure experiments ($\leq 5\%$ by volume). Preliminary X-ray measurements on the decompressed perovskite samples showed no evidence for strong preferred orientation (H. R. Wenk, C.M. and R.J., manuscript in preparation).

Our methods for measuring the strength of minerals at high pressures and room temperature have been described elsewhere^{14,18}. To summarize, we consider the momentum balance of the sample in the diamond cell to obtain a relation for the shear stress τ in terms of the maximum pressure gradient $\partial P/\partial r$ and the sample thickness h : $\tau = \frac{1}{2}h \partial P/\partial r$ ($\partial P/\partial r$ is measured perpendicular to the loading axis). We measured the pressure gradients at the maximum compression after the conversion to the high-pressure phases, and then on decreasing pressure. The sample thicknesses were measured at zero pressure and then corrected to the values at high pressure using the room-temperature equations of state^{19,20}.

For a number of reasons we infer that the deformation in these high-pressure experiments occurs by dislocation glide. First, low-temperature deformation experiments on a range of silicates and oxides have demonstrated that dislocation glide can be the predominant deformation mechanism, even at low pressures. This has been found in hardness measurements on olivine³, and MgSiO_3 perovskite²¹, and in creep experiments on $\text{MgO}(\text{Al}_2\text{O}_3)_{1.1}$ spinel at $T/T_{\text{melting}} < 0.3$ (ref. 22). Second, the confining pressures in our experiments are significantly higher than both the strength of the samples and the brittle–ductile transition pressure for silicates at room temperature ($\leq 7\text{ GPa}$;

* Now at: Geophysical Laboratory, 5251 Broad Branch Road NW, Washington, DC 20015-1305, USA.

ref. 23). Finally, there is no evidence for fracturing or acoustic emissions associated with the deformation at the pressures of the strength measurements²⁴.

Figure 1 shows our results for olivine, spinel, perovskite and perovskite+magnesiowüstite assemblages to average pressures of 60 GPa at room temperature. A general comparison of the data shows that there are dramatic differences between the strengths of olivine and of the high-pressure silicate and oxide phases. Also the strength of olivine increases sharply with compression. When these measurements are extrapolated back to ambient conditions, they are in good agreement with estimates of the strength derived from hardness measurements at room temperature³ (Fig. 1).

The high-pressure phases are significantly stronger than olivine, although the pressure dependence of their strength is much lower. Whereas the strength of olivine is 2.7 GPa at room temperature and pressure³, our measurements indicate the strengths of perovskite and spinel are ~ 6.0 and ~ 7.9 GPa, respectively, at ambient conditions. The difference in strength between olivine and γ -spinel is consistent with the qualitative estimates of Sung *et al.*²⁵. Also, our data agree well with the recent hardness (H) measurements of 18 ± 2 GPa on single crystals of MgSiO_3 perovskite at ambient conditions²¹. Converting these hardness data to values of the polycrystalline strength S is not straightforward because it requires information about the constraint factor C in the experiment³ ($H \approx 2SC$). Taking the maximum range of values for C ($1.1 < C < 3.0$; ref. 3), the hardness measurements are consistent with a strength of 5.9 ± 3.2 GPa.

We find that there is no significant difference between the high-pressure strength of perovskite and of the perovskite+magnesiowüstite assemblage formed from San Carlos olivine. Because the strength of the composite should obey the standard rules of mixing for the two phases²⁶, the data suggest that (Mg, Fe)O and silicate perovskite have similar strengths at room temperature and that magnesiowüstite is at least 50% stronger than MgO at high pressures and room temperature¹⁴. This last result is in quantitative agreement with the increase in strength

that is observed with increasing iron concentration in the solid solution $\text{Mg}_{1-x}\text{Fe}_x\text{O}$ (ref. 27).

Our observation that γ -spinel is stronger than perovskite between 0 and 60 GPa at room temperature is consistent with previous measurements of a decrease in strength associated with increases in atomic coordination and nearest-neighbour distances across high-pressure phase transformations^{28,29}. When compared to low-pressure experiments, the data in Fig. 1 indicate that the transformation to octahedrally coordinated silicon significantly influences the rheology of mantle silicates. For example, the relative strengths of perovskite and γ -spinel are the opposite of the corresponding low-pressure phases (enstatite is stronger than olivine³⁰). Moreover, their strengths are counter to the trend for tetrahedrally coordinated silicates in which increasing silica content and polymerization increases the strength³¹. In general, our results are consistent with the predictions of previous molecular-dynamics calculations for silicate melts³², and they have important implications for creep models of the mantle because they show that changes in rheology and elastic properties may have opposite signs across high-pressure phase transitions (compare ref. 33).

Electron-microscope observations³⁴ have shown that shocked samples of $(\text{Mg}_{0.76}\text{Fe}_{0.24})_2\text{SiO}_4$ γ -spinel have the same dislocation structures and high-temperature deformation mechanisms as MgAl_2O_4 spinel, a structurally similar mineral which is significantly stronger than olivine. In both spinels, dislocations on the primary glide systems ($\{110\}\{110\}$, $\{110\}\{111\}$) dissociate into partial dislocations on planes perpendicular to the glide plane^{34,35}. This process gives MgAl_2O_4 spinel its high strength because the partial dislocations must climb out of their plane for slip to occur³⁵. In this way, the dislocation splitting inhibits further deformation and produces a significant increase in strength of MgAl_2O_4 spinel relative to olivine with increasing temperature (from a factor of two at 300 K to more than twenty above 1,600 K (refs 1–5, 36)).

Because of the similar structures and deformation mechanisms between silicate and Mg–Al spinel, we use the scaling results of Frost and Ashby³⁷ and Karato³⁸ to estimate the high-temperature strength of γ -spinel from our room-temperature data. The previous work shows that one can describe the creep properties of all spinel structures with a single relation when the shear stress and temperature are normalized by the shear modulus and melting temperature for the particular mineral (for example, τ/μ , T/T_m). This has important implications for our results because silicate and Mg–Al spinel have comparable shear moduli and melting temperatures (within 10%; refs 39, 40), indicating that the temperature dependence of the flow stress ($\partial \ln \tau / \partial (1/T)$) should be similar for the two minerals, and that the difference in strength between γ -spinel and olivine should increase significantly with temperature. Presumably, these results for γ -spinel extend to the modified spinel structure β - Mg_2SiO_4 as well. For comparison, we note that low-pressure creep experiments⁴¹ also show that silicate garnets (almandine and grossular) are significantly stronger than olivine up to $\sim 1,000$ K. These results are consistent with observations of mantle xenoliths that show relatively pristine, undeformed garnets coexisting with highly strained olivine and enstatite⁴².

Because spinels (both β - and γ -phases) and garnets compose the majority of the minerals at depths between 400 and 650 km (~ 60 and 15 – 30% , respectively; ref. 43), our results suggest that transition zone may be significantly stronger than the overlying upper mantle and that it could be stronger than the lower mantle if spinel retains its high strength relative to perovskite at high temperatures. Recently, Karato³⁸ reached a similar conclusion based on semi-quantitative predictions for the strength of garnets at high temperatures. Our data provide experimental support for the hypothesis of a strong transition zone. Our conclusions are based on the properties of silicate spinels, the dominant minerals of the transition zone, rather than garnet, which is a minor component and which may have little effect on the bulk

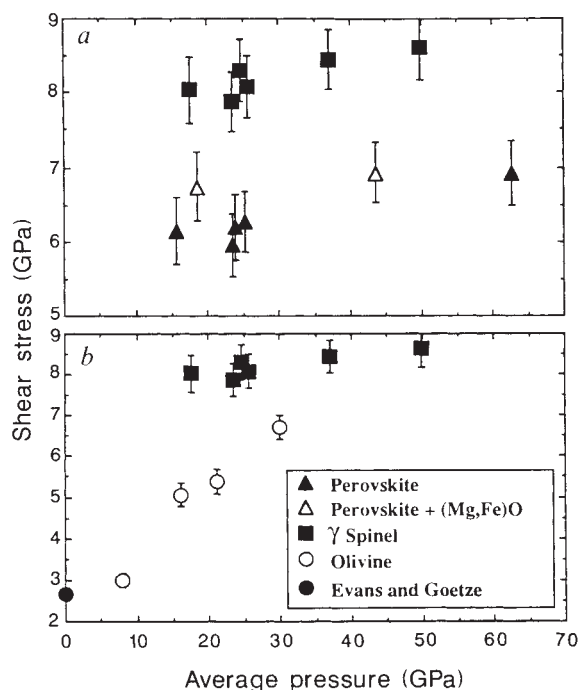


FIG. 1 a, Room-temperature strength of γ -spinel, $(\text{Mg}_{0.88}\text{Fe}_{0.12})_2\text{SiO}_4$ perovskite and perovskite + magnesiowüstite assemblages $(\text{Mg}_{0.88}\text{Fe}_{0.12})_2\text{SiO}_4$ bulk composition) with increasing pressure. b, Room-temperature strength of $(\text{Mg}_{0.88}\text{Fe}_{0.12})_2\text{SiO}_4$ olivine and $(\text{Mg}_{0.88}\text{Fe}_{0.12})_2\text{SiO}_4$ γ -spinel with increasing pressure. For comparison, we show the strength derived from hardness measurements on olivine at room temperature and pressure³.

rheological properties. Of course, the largest uncertainty for our and for previous proposals are the large extrapolations in strain rate required of all deformation studies, because mantle strain rates are at least five orders of magnitude below the values that can be achieved in the laboratory.

At present, observational tests of the laboratory results may also be difficult, because a strong rheological layering between the upper mantle and transition zone would not be well resolved in the analysis of post-glacial uplift data. To illustrate this point, we considered the glacial loading of an Earth model with an infinitely viscous transition zone. For this purpose, we constructed a two-dimensional layered model in which the upper and lower mantle have the same viscosity and the transition zone is a purely elastic layer. We assumed that the layers are distinct and that they have average densities and elastic properties corresponding to the values for the upper mantle, transition zone and lower mantle. We also assumed that there is no flow between the layers (this is reasonable given the small strains associated with post-glacial uplift). Applying a static force balance to the elastic layer, we then calculated the degree of compensation (σ/σ_0) for stresses of different wavelengths that are exerted on top of the transition zone⁴⁴ (Fig. 2).

Although this is a simple and extreme model, with a non-ductile transition zone, the results have important implications for the resolution of viscous layering in the mantle: for the predominant length scales of the stresses associated with continental glaciers ($4,300 < \lambda < 9,000$ km; $\sigma/\sigma_0 \approx 1$), the uplift samples the viscosity of the lower mantle because these long-range stresses are transmitted through the transition zone by elastic deflection. That is, even if its viscosity is unrealistically high, the transition zone might be 'transparent' in the analysis of continental-scale post-glacial uplift and of the time derivative of the low-order gravity potential harmonics (J_2, J_3). Also, at much shorter length scales, the uplift does not probe the rheology of the transition zone because the thickness of the overlying upper mantle is comparable to the scale of the deformation (for example, Lake Bonneville and Greenland, $\lambda \approx 500$ km, ref. 8; $\sigma/\sigma_0 \approx 0$). At the intermediate scales ($1,500 < \lambda < 4,000$ km), however, the effects of a stiff transition zone would be significant.

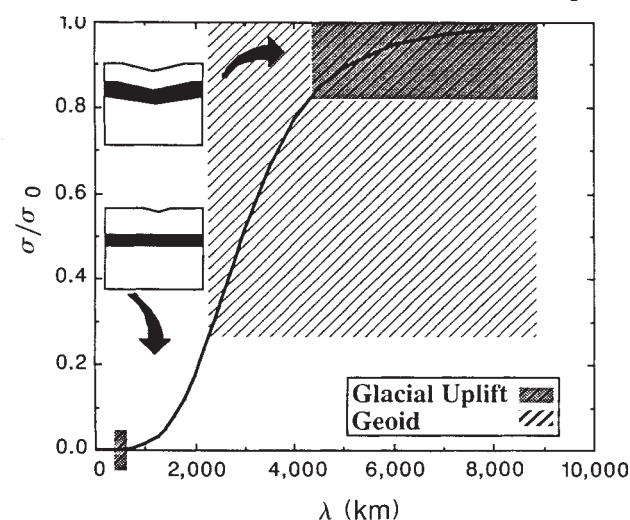


FIG. 2 Degree of compensation for the loading of a 250-km-thick layer with the elastic properties and density of the transition zone. Above and below the layer are fluids with the densities of the upper and lower mantle. For values of $\sigma/\sigma_0=1$, all of the stresses exerted on top of the layer are transmitted to the base by elastic deflection. Alternatively, for $\sigma/\sigma_0=0$, all of the stress is supported by the layer, and there is no deflection (see insets). For comparison, we show the predominant length scales of the stresses corresponding to the loads that have been analysed by models of post-glacial uplift⁶⁻¹² (Fennoscandian, North American, Lake Bonneville and Greenland glaciers) and the geoid (all glaciers and subduction zones). At the depth of the transition zone, the predominant length scale of the stresses is 5.2 times the radius of the load on the surface of the Earth⁸.)

Of course, such a rheological response might be obscured in post-glacial uplift data because the interpretation of short and intermediate wavelengths is complicated by the possibility of lateral variations in both the ice-load history and viscosity of the mantle lithosphere.

If the viscosity of the transition zone is significantly higher than that of the upper or lower mantle, our calculations also suggest that there could be discrepancies between the mantle viscosity inferred from observations of post-glacial uplift and from modelling of the geoid because of the different length scales corresponding to the boundary conditions for these two sets of observations; indeed, such discrepancies exist^{8-11,13} (Fig. 2). Whereas the data of glacial uplift are skewed between long and short dimensions, the scales for the boundary conditions in geoid models (glacial loading and subduction zones) cover a range of dimensions for which the rheology of the transition zone could be important (Fig. 2). In general, these conclusions are supported by recent geoid models that admit a high-viscosity transition zone as a possible solution⁴⁵ and by models that document the effect of a strong transition zone on the geoid and post-glacial rebound⁴⁶. Presumably, more precise measurements of post-glacial uplift, and of the geoid and its time-dependence, will provide important constraints for resolving the viscosity of the transition zone and the differences between the models^{47,48}. At present, there is considerable motivation for this work because a sharp increase in viscosity at 400 km could significantly influence the timescales for mantle convection and the dynamics of slab subduction into the transition zone. □

Received 31 July; accepted 28 September 1990.

1. Raleigh, C. B. *J. geophys. Res.* **73**, 5391-5406 (1968).
2. Goetze, C. *Phil. Trans. R. Soc. A* **288**, 99-119 (1978).
3. Evans, B. & Goetze, C. *J. geophys. Res.* **84**, 5505-5524 (1979).
4. Kohlstedt, D. L., Nichols, H. P. K. & Hornack, P. *J. geophys. Res.* **85**, 3122-3130 (1980).
5. Karato, S., Paterson, M. S. & Fitzgerald, J. D. *J. geophys. Res.* **91**, 8151-8176 (1986).
6. Haskell, N. A. *Am. J. Sci.* **33**, 22-28 (1937).
7. McConnell, R. K. *J. geophys. Res.* **73**, 7089-7105 (1968).
8. Cathles, L. M. *The Viscosity of the Earth's Mantle* (Princeton University Press, 1975).
9. Peltier, W. R. *A. Rev. Earth planet. Sci.* **9**, 199-225 (1981).
10. O'Connell, R. J. *Geophys. J. R. astr. Soc.* **23**, 299-327 (1971).
11. Yuen, D. A., Sabadini, R. C., Gasperini, P. & Boschi, E. *J. geophys. Res.* **91**, 11420-11438 (1986).
12. Peltier, W. R. *Nature* **304**, 434-436 (1983).
13. Hager, B. *J. geophys. Res.* **89**, 6003-6013 (1984).
14. Meade, C. & Jeanloz, R. *J. geophys. Res.* **93**, 3261-3269 (1988).
15. Jeanloz, R. et al. *Science* **197**, 457-459 (1977).
16. Heinz, D. L. & Jeanloz, R. *J. geophys. Res.* **92**, 11437-11444 (1987).
17. Mao, H. K., Bell, P. M., Shaner, J. W. & Steinberg, J. *J. appl. Phys.* **49**, 3276-3283 (1978).
18. Meade, C. & Jeanloz, R. *Phys. Rev. B* **42**, 2532-2535 (1990).
19. Jeanloz, R. & Thompson, A. B. *Rev. Geophys. Space Phys.* **21**, 51-74 (1983).
20. Knittle, E. & Jeanloz, R. *Science* **235**, 669-670 (1987).
21. Karato, S., Fujino, K. & Ito, E. *Geophys. Res. Lett.* **17**, 13-16 (1990).
22. Kirby, S. & Veyssière, P. *Phil. Mag.* **A41**, 129-136 (1980).
23. Paterson, M. S. *Experimental Rock Deformation* (Springer, New York, 1978).
24. Meade, C. & Jeanloz, R. *Nature* **339**, 616-618 (1989).
25. Sung, C. M., Goetze, C. & Mao, H. K. *Rev. sci. Instrum.* **48**, 1386-1391 (1977).
26. Kelly, A. & Tyson, W. R. in *High Strength Materials* (ed. Zackay, V. F.) 578-602 (Wiley, New York, 1965).
27. Srinivasan, M. & Stoebe, T. G. *J. appl. Phys.* **41**, 3726-3730 (1970).
28. Meade, C. & Jeanloz, R. *Science* **241**, 1072-1074 (1988).
29. Meade, C. & Jeanloz, R. *J. geophys. Res.* **93**, 3270-3274 (1988).
30. Ross, J. V. & Nielsen, K. C. *Tectonophysics* **44**, 233-261 (1978).
31. Brainer, S. *Relaxation in Viscous Liquids and Glasses* (American Ceramic Society, Columbus, 1985).
32. Angell, C. A., Cheeseman, P. & Tamaddon, S. *Science* **218**, 885-887 (1982).
33. Sammis, C. G., Smith, J. C., Schubert, G. & Yuen, D. A. *J. geophys. Res.* **82**, 3747-3761 (1977).
34. Madon, M. & Poirier, J. P. *Science* **207**, 66-68 (1980).
35. Doukhan, N., Duclos, R. & Escaig, B. *J. Phys.* **40**, 381-387 (1979).
36. Brethau, T., Castaing, J., Rabier, J. & Veyssière, P. *Adv. Phys.* **28**, 835-1014 (1979).
37. Frost, H. J. & Ashby, M. F. *Deformation Mechanisms Maps* (Pergamon, Oxford, 1982).
38. Karato, S. *Phys. Earth planet. Inter.* **55**, 234-240 (1989).
39. Weidner, D. J., Sawamoto, H., Sasaki, S. & Kumazawa, M. *J. geophys. Res.* **89**, 7852-7860 (1984).
40. Ito, E. & Takahashi, E. *Nature* **328**, 514-517 (1987).
41. Smith, B. K. *thesis*, Univ. of California, Berkeley (1982).
42. Nicolas, A. & Poirier, J. P. *Crystalline Plasticity and Solid State Flow in Metamorphic Rocks* (Wiley, London, 1976).
43. Weidner, D. J. *Geophys. Res. Lett.* **12**, 417-420 (1985).
44. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
45. Ricard, Y., Vigny, C. & Froidevaux, C. *J. geophys. Res.* **94**, 13739-13754 (1989).
46. Hong, H. J., Yuen, D. A. & Wu, J. *Phys. Earth planet. Inter.* (in the press).
47. Mitrovica, J. X. & Peltier, W. R. *J. geophys. Res.* **94**, 13651-13671 (1989).
48. James, T. S. & Morgan, W. J. *Geophys. Res. Lett.* **17**, 957-960 (1990).

ACKNOWLEDGEMENTS. We thank T. Ahrens, D. Anderson, L. Cathles, U. Christensen, B. Hager, R. Hazen, R. Hemley, S. Karato, S. Morris, B. O'Neill, M. Richards, S. Sacks, P. Silver and D. Yuen for comments and discussions. This work was supported by the NSF.

Azimuthal anisotropy of velocity in the mantle lid beneath the Basin and Range province

Noureddine Beghoul & Muawia Barazangi

Institute for the Study of the Continents and Department of Geological Sciences, Cornell University, Ithaca, New York 14853, USA

THE Basin and Range province in western North America is one of the largest and best-studied regions of Cenozoic continental extension^{1,2}, yet there is still considerable debate surrounding the geological history of the province, and the cause of the observed widespread rifting. Although recent geological and crustal geophysical observations² constrain crustal-scale processes within the Basin and Range, more information is needed on the structure and rheology of the mantle part of the lithosphere (the mantle lid), to distinguish between competing geodynamic models^{1,2}. Here we present seismic data showing an azimuthal variation of $\sim 3.2\%$ in the P-wave velocity in the shallowest mantle beneath the Basin and Range. The direction of maximum seismic velocity coincides with the NW–SE direction of Neogene and present-day tectonic extension, suggesting that penetrative deformation associated with this extension has induced a preferred orientation of the minerals in the mantle lid. Our results thus imply that deformation associated with present-day surface rifting extends at least 30 km into the underlying mantle.

We used the International Seismological Centre (ISC) database and the two-station method³ to search for azimuthal anisotropy beneath the Basin and Range Province. The paths between the station pairs are restricted to this province (Fig. 1). Because the quality of the ISC data is variable at regional distances, we have developed a method for selecting a data set with reduced errors for P_n velocity measurements³. (P_n is a high-frequency P-wave travelling only in the mantle lid.) The depth of sampling of this wave in the Basin and Range is restricted to the relatively thin mantle lid (~ 30 km thick) beneath the Moho^{4,5}.

The different constraints used in our data selection are presented in ref. 3. The selected events are located at epicentral distances where P_n is the first arrival ($2\text{--}12^\circ$ for the Basin and

Range), and the source is at or very near the azimuth of the two-station pair. This technique largely eliminates hypocentre mislocation errors and errors due to variations in the crustal structure near the source, because only the difference in travel times at the two stations is used. The error on the measured velocities is estimated to be about $\pm 0.05 \text{ km s}^{-1}$.

Figure 1 shows all (110) combinations of pairs of stations beneath which P_n has been calculated. Based on 254 determinations, the average P_n velocity is 7.85 km s^{-1} . We have carefully documented that observed variations in the 254 determinations are not simply related to epicentral distances and/or to the distance between station pairs. This convinced us that the velocity gradient in the mantle lid beneath the Basin and Range is not significant. This has been confirmed by recent studies⁵, and therefore does not appreciably affect our P_n determinations. Our value for the average P_n velocity is similar to other determinations^{6–8}.

The plot of velocity as a function of azimuth (Fig. 2a) is clearly a sinusoidal function. The function V that describes this distribution of points can be written

$$V = \bar{v} + \delta V(\Phi) \quad (1)$$

where $\bar{v}$ is the mean velocity (7.85 km s^{-1}) of the set of points of Fig. 2a and $\delta V(\Phi)$ is a perturbation around this mean. The best fit to the perturbation $\delta V(\Phi)$ is

$$\delta V(\Phi) = A \cos(\Phi + a) + B \cos(2\Phi + b) \quad (2)$$

where Φ is the azimuth in degrees measured from the north, and A , B , a and b are constants. There is an apparent uniform scatter in the data of $\sim 0.1 \text{ km s}^{-1}$ around the least-squares fit. This scatter increases around certain azimuths, however, and can be $0.25\text{--}0.30 \text{ km s}^{-1}$. There is similar scatter in oceanic data^{9,10}. This scatter cannot be explained by only lateral variations in crustal velocity and/or reading errors but is probably the effect of complex structures in the upper mantle. The fit in Fig. 2a provides a minimum estimate of the variation of velocity with azimuth of $\sim 3.2\%$. Figure 2a also shows that there is no systematic pattern that would indicate a variation of the amount of anisotropy with the average distance to the two stations. At short distances P_n is a head wave whereas at larger distances P_n propagates as a ray with a turning point in the lid⁴, and hence larger distances correspond to greater sampling depth. Thus, there is no apparent variation of P_n velocity with depth in the mantle lid.

The dependence of the velocity on $\cos(\Phi)$ in equation (2) is related to the Moho dip and/or on lateral variations of crustal velocity, whereas the dependence on $\cos(2\Phi)$ could be due to intrinsic anisotropy, Moho curvature, crustal and/or mantle layering, or quadratic variations of lateral velocity in the crust and/or in the mantle¹¹. The magnitude of the constants of A and B quantify the amount of dip (and/or lateral variations of crustal velocity) and anisotropy (and/or Moho curvature), respectively. We obtain the best fit to our data using $A = 0.050 \pm 0.010$ and $B = 0.124 \pm 0.012 \text{ km s}^{-1}$. Some error in these values arises from the non-uniform distribution of the data in Fig. 2a; this is especially so because of the dominant weight given to some sets of points that are concentrated around certain azimuths. The F -test can be used to test the relative statistical significance of the two parameters $A \cos(\Phi + a)$ and $B \cos(2\Phi + b)$. Their values of F are 9.4 and 80.2, respectively. Because the size of the population is relatively large (254 points) and because both values of F are much greater than one, the contributions of $A \cos(\Phi + a)$ and $B \cos(2\Phi + b)$ are significant at the 0.01 level ($F_{0.99}$) and, hence, cannot be ignored in the fit. The contribution of $B \cos(2\Phi + b)$ value is about eight times larger than $A \cos(\Phi + a)$ value.

The azimuths of maximum and minimum velocity are $N65^\circ \pm 10^\circ$ and $N27^\circ \text{ E} \pm 10^\circ$, respectively. The direction of the average Quaternary extension of the interior of the Basin and Range is orientated approximately $N60^\circ \text{ W}$, whereas the

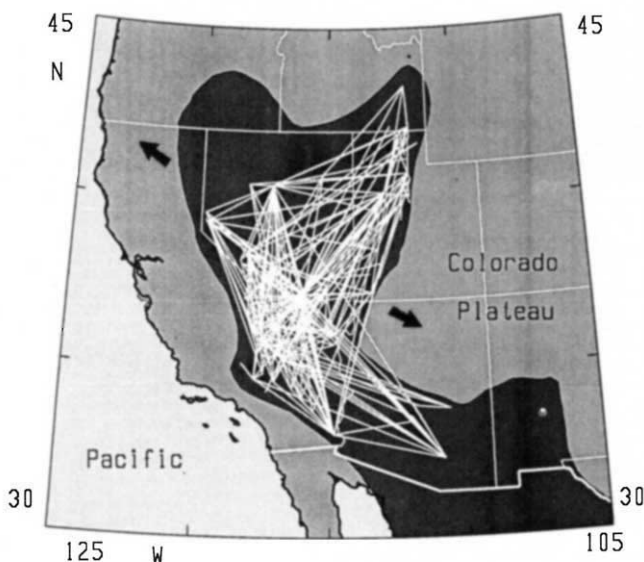


FIG. 1 Map of the area. Dark shaded zone represents approximate outline of the Basin and Range Province. Solid white lines connect pairs of stations beneath which P_n velocity is calculated. The azimuth coverage is satisfactory. Arrows indicate the direction of present-day extension.

Ranges trend is about N25° E^{12,13}. This indicates that the direction of maximum velocity in the mantle lid is parallel (within the precision of the measurement) to the direction of present-day extension. This observation is similar to the observations made in oceanic lithosphere where the direction of highest velocity in the oceanic mantle lid is found to be parallel to the direction of extension^{9,10}.

Moho curvature, mantle layering and quadratic horizontal velocity gradients in the crust and/or mantle can produce the same dependence of velocity on azimuth as intrinsic anisotropy. These alternatives, however, cannot explain our observations. To be explained by Moho curvature alone in the 11 × 7.5° area, $B = 0.124 \text{ km s}^{-1}$ would require a calculated Moho depth¹¹ varying from ~30 km to ≥90 km, which is unrealistic⁷. Nor can subvertical layering explain the observed anisotropy. If one assumes that isotropic evenly spaced dykes penetrate a perfectly isotropic mantle lid and are parallel to the basins and ranges at the surface, then the high velocity should be parallel to the ranges, but this is not observed. Also, the elastic constants for the rheology of the Basin and Range province, calculated from the observed anisotropy, do not fulfil a set of equations postulated by Backus¹¹ for any subhorizontal fine-layered medium. Thus the observed anisotropy is probably not related to lamination (subhorizontal discontinuous intrusions) in the uppermost mantle. The B value of 0.124 km s^{-1} would require a lateral quadratic variation of 30% in the mean crustal velocity; in contrast, the observed⁷ maximum variation is only ~4%. Moreover, no lateral quadratic variation in the velocity of the mantle lid is observed. We therefore conclude that the observed velocity variation must be due to intrinsic anisotropy.

The observed anisotropy of P_n (a minimum of 3.2% of the average velocity) is ~2.5 times smaller than that in the oceanic lithosphere^{9,10}, and about half that found in the continental Rhine graben region^{14,15}. Bamford *et al.*¹⁶ report up to 3% azimuthal anisotropy across the western United States including the Sierra Nevada, the Basin and Range and the Colorado plateau. The direction of fastest velocity is about N70° E ± 10°. Errors in these studies may be large, however, as the refraction database used varies considerably in quality and because the data of Bamford *et al.* sample several major tectonic provinces with different velocity structures in the uppermost mantle. A recent study¹⁷ documents ~4% azimuthal anisotropy in the mantle lid beneath the Rio Grande rift with a direction of the highest velocity trending about east-west, that is, approximately parallel to the direction of extension. A similar magnitude of anisotropy was reported in recent observations of shear-wave splitting in the Basin and Range Province in Nevada¹⁸, with an average direction of fastest polarization orientated ENE-WSW, although the direction of present-day extension is NW-SE. The apparent discrepancy with our results may be explained by the fact that the results of Savage *et al.*¹⁸ apply to about the upper 100 km of the mantle, whereas our velocities are restricted to the 30-km-thick mantle lid beneath the Moho. Thus, the results of the two experiments may not be in conflict because the measured anisotropies represent signatures of different parts of the upper mantle. A passive PASSCAL experiment using shear-wave splitting of the phase PmS (converted from a compressional to a shear-wave at the Moho) shows the same direction of anisotropy in the lower crust as we found for the mantle lid¹⁹.

In the following we show that the measured anisotropy can be interpreted as being due to the Cenozoic extensional event and not the result of the cumulative signature of earlier tectonic phases that the Basin and Range has experienced since Precambrian time. We used Ribe's continuum model for lattice preferred orientation²⁰ to model the observed signal of anisotropy and to evaluate the amount of extension necessary to reproduce the observed variation of velocity with azimuth. We assumed that: (1) The anisotropy of the mantle lid is mainly controlled by olivine. This is a reasonable assumption because ~66% of the lid material is made of olivine, and olivine anisotropy in the

horizontal plane is 15%²¹⁻²³. In contrast, the orthopyroxenes form only 20% of the total lid material with only 10% anisotropy in the horizontal plane²³. (2) The volume of partial melt injected in the mantle lid was negligible relative to the total volume of the mantle lid and hence that the volume of mantle-lid material was conserved in pre- and post-Neogene extension. (3) The mantle lid was isotropic (random distribution of fast axes) when the last major Neogene extension started. (4) the seismic velocity for a given time t and a given azimuth Φ (with ±10° error on the azimuth) is linearly dependent on the relative concentration of the fast and the slower axes of the horizontal plane in the interval $[\Phi - 10^\circ, \Phi + 10^\circ]$. The effect of the vertical mineral axes is neglected because it is mostly cancelled by the use of the two-stations method.

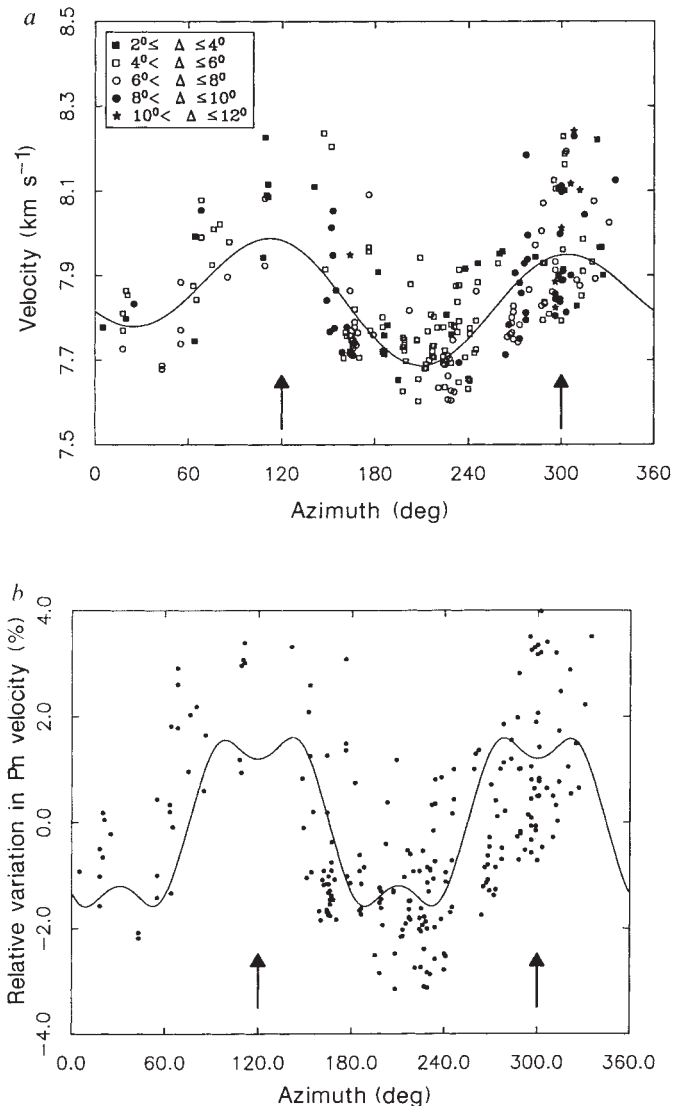


FIG. 2 a, Plot of calculated P_n velocity against azimuth (measured from north). Different symbols correspond to given ranges of epicentral distances. The distance of the event correlates with depth of penetration of the turning point of the P_n ray. There is no systematic pattern that would suggest any dependence of amount of anisotropy on distance, and hence on depth. Best fit to this distribution of points is given by equation (2). B (which is related to anisotropy) is much larger than A (which is related to dip and/or crustal-velocity variations). Arrows indicate the approximate azimuth of present-day extension. b, The solid line represents the relative variation of P_n velocity with azimuth that best fits Basin and Range data points (see text for assumptions). This best fit is achieved for 45% of lithospheric extension (pure shear). The variation of P_n velocity is ~4% for 45% extension and the largest variation in the velocity occurs in the direction of present-day extension indicated by the arrows.

Using these assumptions, the best least-squares fit to the data is obtained for ~45% of lithospheric extension (Fig. 2b), and leads to a relative variation in P_n velocity of ~4%, which is similar to the observed anisotropy. As discussed earlier, the curve in Fig. 2b provides only minimum estimates of the anisotropy of P_n velocity and lithospheric extension because of the uneven distribution of data. It follows that, besides the alignment of the fast direction of seismic velocity with the direction of present-day extension, the magnitude of the extension measured from the observed anisotropy (~45%) is within the range of the amount of Cenozoic extension measured from surface geology^{24,25}. These observations favour the inter-

pretation that the observed anisotropy is due to Cenozoic extension of the lithosphere rather than to the cumulative signature of the earlier tectonic phases. Thus, only a few million years have been sufficient to reset the old lattice preferred orientation of the mantle-lid minerals and record the signature of the present tectonic regime. Clearly, many simplifying assumptions have been made to reach this conclusion and it may therefore not be unique. Our observations strongly suggest, however, that the pervasive surface rifting of the Basin and Range is not only a crustal process but also affects mantle-lid deformation. □

Received 28 February; accepted 16 October 1990.

1. Smith, R. B. & Eaton, G. P. (eds) *Mem. geol. Soc. Am.* **152** (1978).
2. Pakiser, L. C. & Mooney, W. D. (eds) *Mem. geol. Soc. Am.* **172** (1989).
3. Beghoul, N. & Barazangi, M. *J. geophys. Res.* **94**, 7083–7104 (1989).
4. Burdick, L. J. *Eos* **70**, 400 (1989).
5. Burdick, L. J. & Helmberger, D. V. *J. geophys. Res.* **83**, 1699–1712 (1978).
6. Bralle, L. W., Hinze, W. J., Frese, von R. R. B. & Keller, G. R. *Mem. geol. Soc. Am.* **172**, 655–680 (1989).
7. Pakiser, L. C. *Mem. geol. Soc. Am.* **172**, 235–247 (1989).
8. Iyer, H. M. & Hitchcock, T. *Mem. geol. Soc. Am.* **172**, 681–710 (1989).
9. Morris, G. B., Raitt, R. W. & Shor, G. G. *J. geophys. Res.* **74**, 4300–4316 (1969).
10. Raitt, R. W., Shor, G. G. & Kirk, H. K. *Tectonophysics* **12**, 173–186 (1971).
11. Backus, G. E. *J. geophys. Res.* **70**, 3429–3440 (1965).
12. Minster, J. B. & Jordan, T. H. *J. geophys. Res.* **92**, 4798–4808 (1987).
13. Zoback, M. L. *J. geophys. Res.* **94**, 7105–7128 (1989).
14. Bamford, D. *Geophys. J. R. astr. Soc.* **49**, 29–48 (1977).
15. Fuchs, K. *Phys. Earth planet. Inter.* **31**, 93–118 (1983).

16. Bamford, D., Jentsch, M. & Prodehl, C. *Geophys. J. R. astr. Soc.* **57**, 397–429 (1979).
17. Singer, P. J. & Sanford, A. R. *Seism. Res. Lett.* **60**, 17 (1989).
18. Savage, M. K., Silver, P. G. & Meyer, R. P. *Geophys. Res. Lett.* **17**, 21–24 (1990).
19. McNamara, D. E., Owens, T. J., Zandt, G. & Randall, G. E. *Geol. Soc. Am. abstract with programs* **21**, A321 (1989).
20. Ribe, N. M. *J. geophys. Res.* **94**, 4213–4223 (1989).
21. Christensen, N. I. & Salisbury, M. H. *J. geophys. Res.* **84**, 4601–4610 (1979).
22. Nicolas, A. & Christensen, N. I. in *Structure and Dynamics of the Lithosphere-Asthenosphere System* Vol 16, (eds Fuchs, K. & Froidevaux, C.) 111–123 (Am. geophys. Un., Washington, DC, 1987).
23. Anderson, D. L. in *Theory of the Earth* 303–335 (Blackwell Scientific, Boston, 1989).
24. Gans, P. B. *Tectonics* **6**, 1–12 (1987).
25. Levy, M. & Christie-Blick, N. *Science* **245**, 1454–1462 (1989).

ACKNOWLEDGEMENTS. We thank C. Stewart, S. Emerman, T. Hearn, B. Isacks, J. Huang, D. Turcotte, A. Hamza and E. Hauser for discussions, B. Isacks, J. Oliver and R. Allmendinger for reviews, and J. Otto, C. Caruso, B. Payne and J. Best for comments. We also thank M. Bevis and J. L. Chatelain for a contouring program. This research was supported by the NSF.

Shared paternity revealed by genetic analysis in cooperatively breeding tropical wrens

Patricia P. Rabenold, Kerry N. Rabenold, Walter H. Piper, Joseph Haydock & Steve W. Zack

Department of Biological Sciences, Purdue University, W. Lafayette, Indiana 47907, USA

POSTPONEMENT of dispersal and breeding to assist in rearing others' young may be favoured if helpers' contributions to the production of close kin exceed their likely reproductive success had they dispersed^{1,2}. Young adult stripe-backed wrens (*Campylorhynchus nuchalis*) remain in natal groups and greatly enhance reproductive success of close kin. Long-term behavioural observations suggest that only the dominant male and female in a group breed³; however, extra-pair parentage has been confirmed biochemically in apparently monogamous birds^{4,5}, including some social species^{6,7}. DNA fingerprinting of wren groups shows that behaviourally dominant males sometimes share paternity with auxiliary males previously thought to be nonreproductive, whereas dominants are the only reproductives among females. Reproduction by auxiliary males (but not females) helps explain the long tenure

of males in helper status and the contrasting combativeness of females in competition for breeding positions outside the natal group^{8,9}, but reproduction by auxiliaries cannot alone explain helping behaviour.

We analysed parentage for 69 offspring produced by 22 social groups. Blood was collected from 260 birds in 1988–89; 110 constituted complete breeding-season membership of 17 social groups that reproduced during those years and 28 additional samples provided tests for 5 other groups in previous years. Histories of group membership and behavioural status were available for the previous 4–12 years.

For each group, genomic blots of *Hae*III and separate *Hinf*I digests were hybridized with Jeffreys' probes 33.6 and 33.15 (ref. 10). *Alu*I blots were also analysed for large families. Five scorers located bands exclusive to each adult on autoradiographs and recorded their presence in offspring lanes. For 53 of 68 offspring, 1 adult female and 1 adult male in the group provided a minimum of 2 exclusive bands (range 2–18, mean 5.9 ± 3.4 for *Hae*III blots with both probes) and the entire offspring banding pattern could be derived only from those two adults. Then, two of us tallied the number of bands for each offspring that were unattributable to each adult combination of potential parents⁵ (intra-group) (Fig. 1). The parents assigned by minimizing unattributable bands in no case contradicted parents assigned by matching exclusive adult bands. The second technique was important for large families with several closely related potential breeders (brothers accumulate because of high natal philopatry in males; Fig. 2) and allowed assignment of parentage for all but one

TABLE 1 Parentage in stripe-backed wrens

Group composition	No. group-years	No. Young	Parentage		
			PM-PF	PF-AM	Outside
2 PRINS only	5	10	9	NA	1
2 PRINS + AM	10	23	19	4	0
2 PRINS + AF	1	2	2	NA	0
2 PRINS + AM and AF	18	34	32	2	0
Totals	34	69	62	6	1

Groups are composed of principal males (PM), principal females (PF) (PRINS for both), auxiliary males (AM) and auxiliary females (AF).

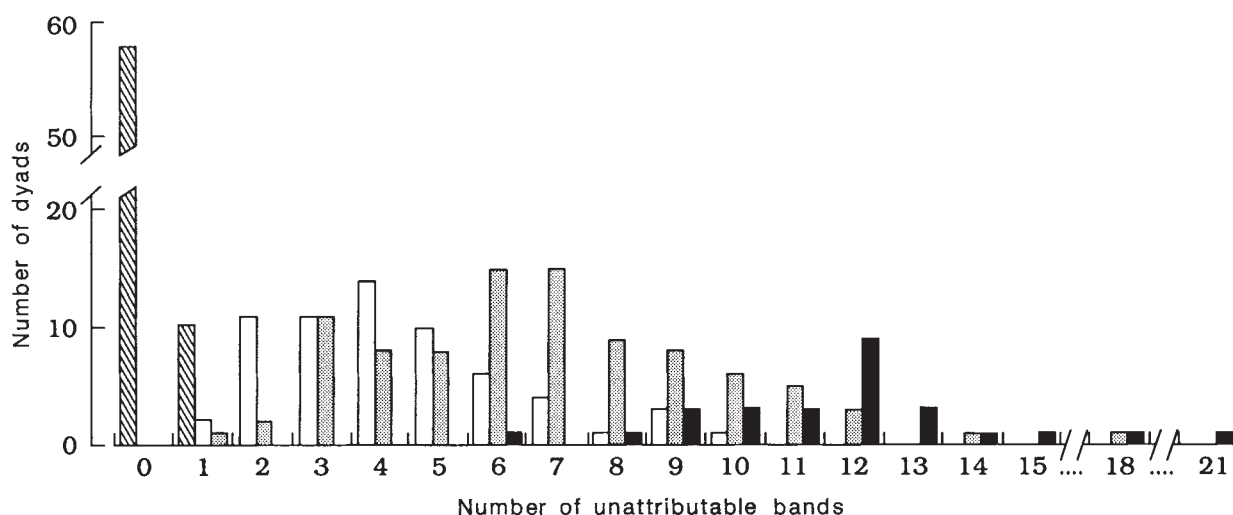


FIG. 1 Distribution of unattributable bands produced by considering every possible adult male-female dyad within groups as potential parents of each offspring, summing across probes 33.6 and 33.15 (ref. 10) with *HaeIII*. Assigned parents (hatched bars) produce an average of 0.15 ± 0.36 unattributable bands in juveniles ($n=68$); combinations of first-order relatives (open bars) excluding parental dyads (for example, mother-brother or brother-sister pairs) produce 4.9 ± 4.0 unattributable bands ($n=64$); first-order relatives with lesser relatives (for example, sister-uncle pairs) (stippled bars) produce 7.4 ± 6.9 ($n=92$); and dyads of less than first-order relatives (for example, uncle-immigrant pairs) (solid bars) produce 11.8 ± 3.3 ($n=22$). Even combinations of first-order relatives are usually distinct from best-fit dyads. Among the 68 offspring with assigned parents, 58 had banding patterns entirely and uniquely derivable from the parental banding patterns. Ten offspring had one novel band unattributable to either assigned

parent; we consider these novel fragments mutations. Occasional unattributable bands in offspring are expected because of high mutation rates in these hypervariable regions of the genome¹⁰, between 0.003 and 0.05 mutations per locus per generation for humans¹² and 0.0035 for house sparrows¹³. We calculate the mutation rate for stripe-backed wrens, based on the occurrence of 10 novel bands in 68 offspring, each of which had on average 54.4 bands across both probes on *HaeIII* blots, as $(10/68)/54.4$ or 0.0027 mutations per locus per generation. The distribution of the best fit (parental assignments) category does not differ from random expectation (Poisson) ($G=1.394$, $p>0.5$). The expected frequency of individuals showing two or more unattributable bands derived from the actual parents is 0.0098. With a probability of <0.01 that actual parents would produce two or more unattributable bands, we call any adult combinations producing two or more unattributable bands in offspring a misassignment.

juvenile (Fig. 3b).

Of the 69 young produced in 34 group-years, 62 (90%) were offspring of the principal male and principal female of their group, and 6 were offspring of the principal female and an auxiliary male. In no case was an auxiliary female identified as the parent of a juvenile and in only one case was extra-group

parentage suggested (Table 1). The 4 reproductive auxiliary males were found in 4 of the 13 groups in which an auxiliary male helped a brother or father and an unrelated principal female who had immigrated during the auxiliary's lifetime. Parentage in three of these four groups was obvious on the basis of matching exclusive adult bands (Fig. 3a). Assignment in the fourth group (Fig. 2) relied on analysis of unattributable bands. In all four groups, paternity was shared with the principal male: 1/2, 2/4, 1/2 and 2/3 juveniles were sired by the auxiliary.

The opportunity for older subordinate males to breed helps to explain the long tenure of auxiliary males in queues for dominant status in their natal groups. Overt challenges to principals are rare, although recent behavioural observations show increased conflict among males in groups with immigrant breeding females (W.H.P., unpublished data). We had previously thought that the genetic benefits of helping behaviour were both indirect (the production of siblings) and direct but delayed

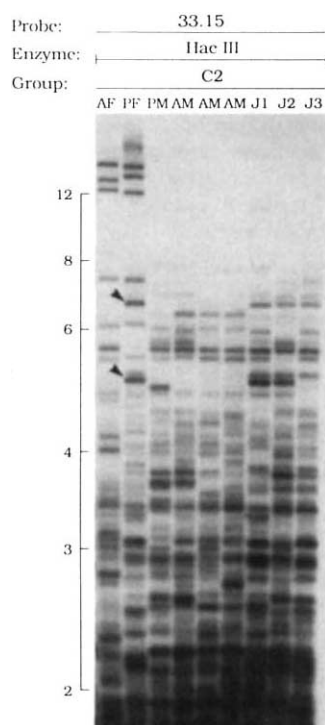


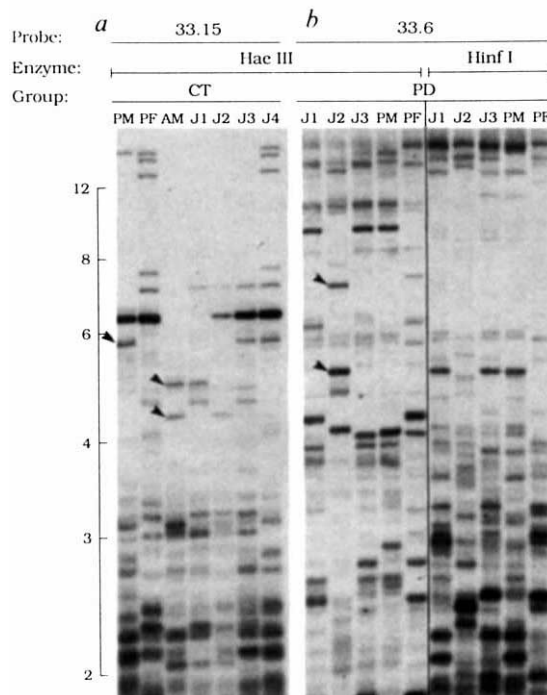
FIG. 2 DNA fingerprints from a large group consisting of two adult females (sexed behaviourally as well as by the presence of female-specific bands above 12 kb; ref. 14), the principal male and three of his male offspring by a previous principal female. The mother is easily assigned to all three juveniles by band-matching; arrows indicate bands important in this assignment. Paternity was tested by pairing the mother with each male in turn and counting the number of unattributable bands produced by each combination for each offspring. For offspring 1 with this probe-enzyme combination, the female in lane 2 (the mother) paired with males in lanes 3, 4, 5 and 6, produced 2, 0, 4 and 5 unattributable bands, respectively. For offspring number 2, the same combinations produced 0, 1, 5 and 6 unattributable bands, and for the third offspring, 2, 0, 7 and 6. The same scoring technique applied to the other five probe-enzyme combinations (using *HinfI* and *AluI*) produced similar results. Using 33.6 with *HaeIII* resolved parentage of offspring 2 and using *AluI* with 33.6 clearly revealed the parentage of offspring 3. The male in lane 4 sired the first and third offspring, whereas the male in lane 3 sired the second. DNA sizes (kb), on the left-hand side.

FIG. 3 DNA fingerprints from two groups of stripe-backed wrens. *a*, (Group CT): The principal male (PM) and principal female (PF) were identified as parents of juveniles 3 and 4, whereas juveniles 1 and 2 were offspring of the auxiliary male (AM) and principal female. Arrows indicate obvious exclusive adult bands confirming this assignment. *b*, (Group PD): Juvenile 2 is the only juvenile of 69 examined not closely related to the rest of its group. Arrows indicate obvious nonmatching bands. Assignment of parents for this juvenile has not been possible among adults in the 10 neighbouring groups. Extra-group parentage is possible, as are juvenile adoption (previously recorded in this population) or accidental capture of a wandering juvenile. DNA sizes (kb), on the left-hand side.

METHODS. Blood samples, stored in phosphate-buffered saline, were incubated overnight at 55 °C after adding SDS to 0.8% and 200 µg Proteinase K, extracted several times in phenol/phenol:chloroform:isoamyl alcohol, then dialysed against TNE₂. DNA (5 µg) were digested with excess restriction endonuclease at 37 °C. Resulting fragments were separated through a 0.8% agarose gel (22 cm) at 20 V for 65 h until all fragments smaller than 1,600 base pairs (bp) were run off, and then transferred to nylon by Southern blot in 10 × SSC buffer. The probes were radiolabelled by primer extension with [³²P]dGTP. Hybridizations were run overnight at 62 °C in 1.5 × SSC, 0.1% SDS, 5 × Denhardt's solution, and 6% w/v polyethylene glycol. Hybridized filters were washed 4 × 30 min at 62 °C in 1.5 × SSC, 0.1% SDS, and exposed to x-ray film at -80 °C. Using *Hae*III with probe 33.15, a mean (f) of 24 scorable bands were produced per lane, 30 with 33.6. The proportion of bands shared (x , calculated as $2e/(2e+a+b)$ where e is the number of fragments of equal mobility and intensity in each pair of adjacent lanes, a is the number of fragments distinct to one of the pair, b the number distinct to the other) among a sample of 24 birds related at a level less than cousins (based on long-term pedigrees assuming monogamy) was 0.27 using 33.15, and 0.26 using 33.6. Assuming independent segregation, the mean population allele frequency (q), where $x=2q-q^2$, is 0.15 for 33.15 and 0.14 for 33.6 with *Hae*III. The expected proportion of bands shared between siblings (s) = $(4+5q-6q^2+q^3)/4(2-q)$ or 0.62 for the wrens¹⁵. In groups containing two adult brothers, the probability of misassigning the uncle as the father (P_u) is s^c , the likelihood that the uncle and father share all exclusively paternal bands c , where $c=(f)[1-(1+q-q^2)/(2-q)]$. For *Hae*III with 33.15, $c=9.3$ and $c=12.2$ with 33.6, so $P_u=0.62^{9.3}$ or 0.012 for 33.15, and 0.003 for 33.6. Using both combinations, the probability that the uncle would share all exclusively paternal bands in an offspring with the actual father is 0.00004. But, there

(improved future fecundity because of the production of young to reciprocate as helpers when the current auxiliary male becomes a principal in that group)³. We have now demonstrated that immediate direct benefits can also accrue. The tendency for older auxiliary males to contribute more to feeding young³ can now be interpreted as a reflection of increasing probability of parenthood, rather than just as investment in future helpers.

Our previous estimate of the indirect component of fitness has been revised downward. If, as in our sample, auxiliary males sire 10% of the offspring in the population, each can expect to sire an average of 0.08 young per yr (demographic data for 1986–1989, $n=212$ auxiliary male-yr). In the subset of groups in which principal females are recent immigrants, auxiliary males sire on average 0.24 young yr⁻¹. The alternative of attempting breeding outside the natal group (unaided pairs are the only option for dispersing males) yields on average 0.26 young yr⁻¹ (1986–1989 data, $n=91$ male-yr). We have previously shown that indirect contributions to genetic representation by increasing sibling production are considerably higher: 0.64 siblings yr⁻¹ (ref. 3). For auxiliary males helping unrelated principal females, we now estimate the indirect component of this enhancement as (0.64 (total)–0.24 (direct)) devalued by half because the remainder are half-siblings, or 0.20 offspring equivalents. Combined preliminary estimates for immediate direct and indirect benefits accruing to such male auxiliaries then yield 0.44 offspring equivalents per yr (full accounting will include delayed effects). Neither the chance of breeding in auxiliary status nor sibling production would be sufficient, even for this subset of auxiliary males, to favour staying in auxiliary status compared to dispersing to attempt breeding; however, these two factors



are often more than two first-order male relatives as potential fathers in a group. The likelihood that an offspring band could be attributed to only one adult male in a group with n first-order male relatives (assuming the true father is included) is $(1-s)^{n-1}$. For wren groups with four first-order male relatives, a paternal band in an offspring lane could be assigned with certainty to a single male with a frequency of 0.055. With $c=21.5$ for both probes, only one exclusive band should match any particular male with an offspring in such a group. Matrices of unattributable bands for each combination of potential parents provide much greater resolution, as mother-uncle dyads should produce $c(1-s)$ or 8.2 unattributable bands on average.

in combination are sufficient to compensate deferred dispersal.

Auxiliary females, unlike males, were never observed to reproduce and apparently derive no direct fitness benefit from helping. The absence of breeding alternatives for females helps explain their investment in fighting over principal positions in non-natal groups with many helpers⁹; they are competing for an opportunity that is valuable in part because it can be monopolized¹¹. The window on genetic relatedness provided by DNA technology allows us to see more clearly the divergence between the sexes, and the importance of ontogeny, in dispersal tendencies and routes to biological fitness. □

Received 30 July; accepted 9 October 1990.

1. Brown, J. L. *Helping and Communal Breeding in Birds* (Princeton University Press, Princeton, New Jersey, 1987).
2. Hamilton, W. D. *J. theor. Biol.* **7**, 1–72 (1964).
3. Rabenold, K. N. *Behav. Ecol. Sociobiol.* **17**, 1–17 (1985).
4. Gowaty, P. A. & Karlin, A. A. *Behav. Ecol. Sociobiol.* **15**, 91–95 (1984).
5. Westneat, D. F. *Behav. Ecol. Sociobiol.* **27**, 67–76 (1990).
6. Wrege, P. H. & Emlen, S. T. *Behav. Ecol. Sociobiol.* **20**, 153–160 (1987).
7. Burke, T., Davies, N. B., Bruford, M. W. & Hatchwell, B. J. *Nature* **338**, 249–251 (1989).
8. Rabenold, K. N. in *Cooperative Breeding in Birds: Long-term Studies of Ecology and Behaviour* (eds Stacey, P. B. & Koenig, W. D.) 157–196 (Cambridge University Press, London, 1990).
9. Zack, S. W. & Rabenold, K. N. *Anim. Behav.* **38**, 235–247 (1989).
10. Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **314**, 67–73 (1985).
11. Grafen, A. *Anim. Behav.* **35**, 462–467 (1987).
12. Jeffreys, A. J., Royle, N. J., Wilson, V. & Wong, Z. *Nature* **332**, 278–281 (1988).
13. Burke, T. & Bruford, M. W. *Nature* **327**, 149–152 (1987).
14. Rabenold, P. P., Piper, W. H., Decker, M. D. & Minchella, D. J. *Genome* (in the press).
15. Jeffreys, A. J., Brookfield, J. F. Y. & Semeonoff, R. *Nature* **317**, 818–819 (1985).

ACKNOWLEDGEMENTS. We thank T. Blohm for providing the field station at Hato Masaguaral and D. Minchella for help with the laboratory work. J. Bennetzen, J. Brown, S. Creel, M. Decker, D. Minchella, R. Howard, J. Lucas, P. Waser and D. Westneat made contributions to this and related papers. The National Science Foundation supported this research.

Sex ratio and habitat limitation promote delayed dispersal in superb fairy-wrens

S. G. Pruett-Jones & M. J. Lewis*

Department of Ecology and Evolution, University of Chicago,
5630 So. Ingleside, Chicago, Illinois 60637, USA

* Department of Zoology, Australian National University, GPO Box 4,
Canberra, ACT 2601, Australia

IN a small proportion of bird species (<3%), offspring delay their dispersal beyond the time when they are reproductively capable, remain on their natal territory, and usually assist the resident breeders, normally their parents, to raise other young¹⁻³. This pattern of social organization and care-giving behaviour, most frequently termed cooperative breeding, is controversial with respect to why offspring delay their dispersal²⁻⁸. In this study we investigate the interaction between habitat limitation, habitat quality, and sex ratio in influencing dispersal decisions in a population of superb fairy-wrens (*Malurus cyaneus*). We show that every non-breeding male (31 of 32; 96.9%) given the opportunity, dispersed and bred. The sex ratio was male-biased and although males did not disperse into vacant but previously occupied territories in the absence of females, the reintroduction of females prompted dispersals. Our data suggest that young males delay their dispersal in response to a limited number of mates and secondarily to habitat limitation.

Superb fairy-wrens are small (9–11 g), insectivorous passerines distributed throughout southeastern Australia⁹. The species exhibits cooperative breeding, and groups typically are composed of one breeding pair with male, but not female, behaviourally subordinate and presumably non-breeding helpers¹⁰ (for simplicity and for lack of genetic data on parentage, see ref. 11, we refer to these individuals as non-breeders). We studied 30 and 54 breeding groups in 1988 and 1989, respectively, at Black Mountain Nature Reserve in Canberra. Group size varied from 2 to 6 adults (1988) and 2 to 9 adults (1989). The adult sex ratio was male biased: during 1988 there were 1.77 males per female and in 1989, 1.85 males per female in the population. Only one group in 1988 and two groups in 1989 had non-breeding females resident on the territory.

During 1989 we conducted manipulations consisting of removing the breeding male from wren groups which, at the time of the experiment, were composed of only two adults of opposite sex. The breeding opportunities (also termed vacancies) created thus consisted of a territory with an unpaired, adult female present. Forty-six removal experiments were conducted on 29 of the experimental territories. In all manipulations, the adult male was captured on his territory, taken to other suitable habitat and released. During the majority ($n = 39$) of removals, the adult female was not disturbed. The other seven were 'removal/release' experiments. Both the male and female were initially removed, the female was held in captivity for two to three days, and then released back onto her territory.

Breeding opportunities were created or occurred naturally for 33 of the 40 non-breeding males (potential dispersers; Table 1). One of the two non-breeders that did not disperse was caring for seven young on his natal territory and during the last week of field work he was given the opportunity to disperse into a territory where a female was incubating eggs. This female would probably not have renested during the 1989 season and the vacancy created may not, in fact, have represented a breeding opportunity. If this male is excluded, only one (3.1%) of 32 potential dispersers did not disperse given a clearly identified breeding opportunity.

All dispersals ($n = 32$; 31 males filled 32 vacancies) were by single individuals from groups either adjacent (28; 87.5%) or

near (4; 12.5%) the territory where the vacancy was created. The mean time to arrival by dispersers (a male vocalizing on and defending the territory) averaged 5.3 hours (day and night hours combined; s.d. = 6.7 h; $n = 24$). When a vacancy was created, the first male to move into the territory and initiate defence was, in every case, the winner and was never displaced. No disperser was rejected by the resident female, and the mean time to acceptance by the female after arrival (indicated by the female and male allopreening or foraging together) was 1.9 hours (s.d. = 1.5 h; $n = 11$).

In the removal/release experiments, no dispersals were recorded after removal of the pair but before the release of the female. In contrast, when females were released back onto their territories, immediate dispersals were recorded.

For 21 of the 29 groups where vacancies were created, dispersals were the outcome observed (S.G.P.-J., manuscript in preparation). For each of these 21 groups, and for each group that sent out dispersers into them, we ranked the territories as to area (in hectares) and maximum group size during the season. Group size in *M. cyaneus* is known to correlate positively with territory quality¹², and we use ranks of group size here as approximate measures of territory quality. Examining the first dispersal that took place in each territory (as an independent set of observations), we can ask whether dispersers moved into larger and/or presumably better-quality territories. With respect to territory size, of the 21 dispersals, males moved into similarly sized territories in 2 (9.5%) cases, larger territories in 8 (38.1%) cases, and smaller territories in 11 (52.4%) cases. With respect to group size, the dispersers moved into groups of similar rank in 3 (14.3%) cases, groups of greater rank in 0 cases, and groups of lesser rank in 18 (85.7%) cases.

The wren habitat on Black Mountain seemed to be saturated. During 1989 only one new territory was initiated where one had not existed during 1988. Males experimentally introduced onto the study area were not able to settle, except for one individual which discovered a breeding vacancy created during this study. Females released onto the study area separate from the removal/release experiments caused dispersals only in vacant but previously occupied areas.

The results of this study from the perspective of the population of non-breeding males are unequivocal. These males disperse and breed if they can and they move into territories irrespective of territory size and, presumably, habitat quality. We suggest that for *M. cyaneus* on Black Mountain, the pathway leading to delayed dispersal by non-breeding males is both simple and direct. Under conditions of habitat saturation, both a shortage of females and limited habitat appear to interact to limit breeding opportunities and thus promote delayed dispersal. Under condi-

TABLE 1 Disposition of non-breeding males in superb fairy-wren groups

Category	Number
Total number of experimental groups	47
Number of groups with non-breeding males	28
Total number of non-breeders (potential dispersers)	40
Number of potential dispersers	
—removed when the breeding male was removed*	2
—disappeared (and were presumed dead) during the season	3
—for which a breeding vacancy was not created	1
—that could not disperse because the breeding male filled the vacancy	1
Remaining number of potential dispersers capable of dispersing given the opportunity	33
Number of potential dispersers above	
—that dispersed given the opportunity	31
—that did not disperse given the opportunity	2

* During two early experiments, a single non-breeding male was removed at the same time as the breeding male.

tions on non-saturated habitat, however, the shortage of mates seems to be sufficient by itself to lead to the natal philopatry of males. These data support, foremost, the limited-mates^{4,10} and, secondarily, the habitat-saturation^{4,5,13} hypotheses for delayed dispersal and breeding in cooperatively breeding species. □

Received 2 July; accepted 4 October 1990.

1. Skutch, A. F. *Auk* **52**, 257–273 (1935).
2. Emlen, S. T. in *Behavioural Ecology: An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 305–339 (Sinauer, Sunderland, 1984).
3. Brown, J. L. *Helping and Communal Breeding in Birds: Ecology and Evolution* (Princeton University, Princeton, 1987).
4. Brown, J. L. *Am. Zool.* **14**, 63–80 (1974).
5. Koenig, W. D. & Pitelka, F. A. in *Natural Selection and Social Behavior: Recent Research and New Theory* (eds Alexander, R. D. & Tinkle, D. W.) 261–280 (Chiron, New York, 1981).
6. Emlen, S. T. *Am. Nat.* **119**, 29–39 (1982).
7. Stacey, P. B. & Ligon, J. D. *Am. Nat.* **130**, 654–676 (1987).
8. Stacey, P. B. & Ligon, J. D. *Am. Nat.* (in the press).
9. Schodde, R. *The Fairy-Wrens* (Lansdowne, Melbourne, 1982).
10. Rowley, I. *Emu* **64**, 251–297 (1965).
11. Brooker, M. G., Rowley, I., Adams, M. & Baverstock, P. R. *Behav. Ecol. Sociobiol.* **26**, 191–200 (1990).
12. Nias, R. C. *Emu* **84**, 178–180 (1984).
13. Selander, R. K. *Univ. Calif. Publ. Zool.* **74**, 1–224 (1964).

ACKNOWLEDGEMENTS. We thank D. Barber, P. Moore, and D. Watt for field assistance, A. Cockburn, R. Mulder and N. Langmore for advice and other assistance, and S. Arnold, J. Coyne, S. Emlen, W. Koenig, F. Pitelka, M. Pruett-Jones and P. Sniegowski for comments. Supported by the Louis Block Fund of the University of Chicago and by the Department of Zoology, Australian National University.

Long-term potentiation of electrotonic coupling at mixed synapses

Xian-Da Yang*, Henri Korn†
& Donald S. Faber*‡

*Neurobiology Laboratory, Department of Physiology, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214, USA

†Laboratoire de Neurobiologie Cellulaire, Département des Biotechnologies (INSERM U261), Institut Pasteur, 75724 Paris Cedex 15, France

LONG-TERM potentiation^{1,2} of chemical synapses is closely related to memory and learning^{3,4}. Studies of this process have concentrated on chemically mediated excitatory synapses. By contrast, activity-dependent modification of gap junctions, which also widely exist in higher structures such as hippocampus and neocortex⁵, has not been described. Here we report that at mixed synapses between sensory afferents and an identified reticulospinal neuron, the electrotonic coupling potential can be potentiated, as well as the chemically mediated excitatory postsynaptic potential, for a prolonged time period using a stimulation paradigm like that which produces long-term potentiation in hippocampus. The effect on coupling is due to an increase in gap-junctional conductance. Our data indicate that the potentiation of both synaptic components requires an increase in intracellular calcium, involves activation of NMDA (*N*-methyl-D-aspartate) receptors, and is specific to the tetanized pathway.

This study is concerned with the monosynaptic connection between eighth nerve sensory fibres and the Mauthner (M-) cell in goldfish (Fig. 1a). Individual afferents terminate on the M-cell's lateral dendrite as large club-endings which have both gap junctions and chemical synapses with that cell; hence extracellular stimulation of the eighth nerve produces a biphasic excitatory response consisting of a fast electrotonic coupling potential followed by a chemical excitatory postsynaptic potential (e.p.s.p.) (Fig. 1b). As the e.p.s.p. starts with an additional delay of about 0.4 ms (Fig. 1b), changes in its amplitude do not influence measurements of the peak coupling potential. Previous work has indicated that only a small fraction of gap junctional

channels are conducting normally⁶ and that the transmitter for the e.p.s.p. is probably glutamate because the response is abolished by the excitatory amino acid antagonist γ -D-glutamylglycine⁷.

To test whether there is long-term potentiation (LTP) at these contacts, we used three paradigms. In all cases, eighth nerve test stimuli were given at a frequency of 1 Hz to obtain stable control e.p.s.ps. For two series, the control was followed by a 4-min period of tetanization, during which short trains of three pulses at 500 Hz were applied at 2-s intervals, the difference between the two being that the strength of the tetanus was adjusted to be subthreshold for M-cell impulse initiation in one case, and just above threshold in the other. We used this conditioning frequency because auditory stimuli of 200–800 Hz maximally activate the eighth nerve fibres projecting to the M-cell⁸. After tetanization, the test pulses were repeated at 1 Hz. The third type of experiment was a control, in which the test pulses were given throughout without tetanization. M-cell input resistance is low, about 200 K Ω , and cannot be reliably measured with a single microelectrode of 10 M Ω . We therefore monitored it indirectly by measuring antidromic spike height, given that the action potential is conducted passively along the somadendritic membrane⁹.

Suprathreshold trains produced persistent enhancement both in the coupling potential and the chemical e.p.s.p. (Fig. 1c). These effects were specific to the tetanized pathway as there was no change in another e.p.s.p. that was produced by stimulating the optic tectum ($n = 5$; Fig. 1d)¹⁰. Furthermore, the extracellularly recorded presynaptic volley remained constant ($n = 5$; Fig. 1e,f), suggesting that the potentiations were not due to recruitment of additional afferents or to an increase in presynaptic spike amplitude.

The magnitude and time course of these effects depended on the strength of the tetanization. Enhancement of the electrotonic coupling was most evident after the stronger conditioning stimuli (Fig. 2a versus c), whereas both paradigms potentiated the chemical e.p.s.p. (Fig. 2b,d). After subthreshold trains, the e.p.s.p. grew steadily over a period of 30 min to an average peak enhancement of 25% (Fig. 2b). By contrast, the suprathreshold protocol produced rapid and persistent potentiation of both synaptic responses, the increase in coupling (Fig. 2c) averaging 25% (range, 0–62%) and that in the e.p.s.p. 75% (range, 31–144%; Fig. 2d). These effects persisted throughout the recording session, which was 70 min after tetanization in the longest experiment. As antidromic spike amplitudes either did not vary or even decreased slightly ($< 15\%$), these changes were not due to a nonspecific increase in M-cell input resistance.

Induction of LTP in the CA1 region of hippocampus requires an increase in postsynaptic intracellular calcium¹¹, which then presumably triggers long-lasting changes^{11,12}. To test whether this is true of the potentiations reported here, the calcium chelator BAPTA (1,2-bis(*o*-aminophenoxy)ethane-*N,N,N',N'*-tetraacetic acid)¹³ was injected into the M-cell before and during suprathreshold tetanization. Both enhancements were blocked (Fig. 3a,b), a result which is similar to those reported for LTP in hippocampal slices^{11,14}.

Most^{15,16}, but not all¹⁷, reported examples of LTP involve the activation of NMDA receptors, which in turn leads to the requisite calcium influx^{18,19}. Considering that the transmitter at this mixed synapse is probably glutamate and that NMDA receptors are involved⁷, we investigated whether their activation also underlies these phenomena, which turned out to be the case. We used similar protocols while the brain was superfused with a specific NMDA-receptor antagonist, namely 100 μ M CPP ((2-carboxypiperazine-4-yl)propyl-1-phosphonic acid)^{20,21} ($n = 4$). In the presence of antagonist, LTP of the two components was completely blocked (Fig. 4a); after washout with saline, the same tetanizing procedure gave immediate enhancement of both coupling and the chemical e.p.s.p., which, in the example illustrated, decayed over a period of about 20 min to steady-state

‡ To whom correspondence should be addressed.

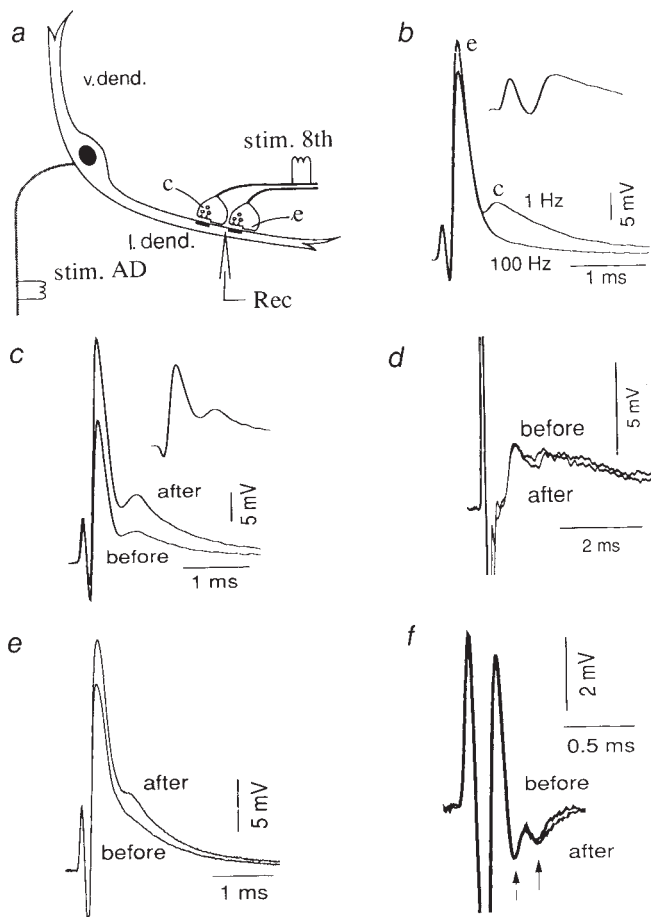


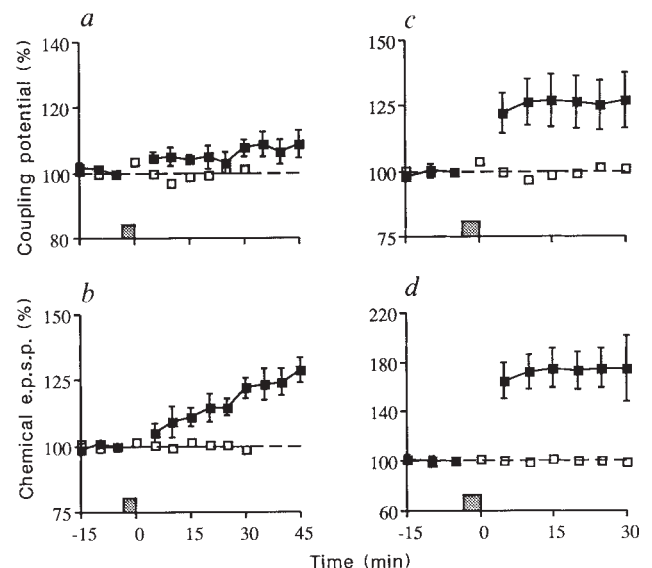
FIG. 1 Supratherapeutic tetanization potentiates both components of the mixed synapses. *a*, Diagram of the experimental arrangement. Note that individual afferents may transmit electrotonically (*e*) and chemically (*c*). *b*, Two superimposed traces of eighth nerve responses obtained with stimulations at 1 Hz and 100 Hz, respectively. The response to 1 Hz stimulation consists of both the electrotonic coupling potential and a chemical e.p.s.p., whereas at 100 Hz, there is only a coupling potential, decreased slightly in amplitude owing to refractoriness of the afferents. The insert is the difference of the two traces showing that the net chemical e.p.s.p. starts after the peaktime of the coupling potential. *c* and *d*, Demonstration of the potentiation and its specificity in one experiment. *c*, Superimposed computer-averaged intracellular recordings of eighth nerve responses obtained 2 min before and 4 min after supratherapeutic tetanization. The insert is the difference between the two, demonstrating that both the coupling potential and chemical e.p.s.p. were enhanced. *d*, In contrast, the corresponding e.p.s.p. caused by stimulating the optic tectum remained the same. *e*, Superimposed eighth nerve responses obtained 3 min before and 6 min after supratherapeutic tetanization. *f*, Corresponding averaged extracellular presynaptic field potentials recorded in the same experiment, with expanded time base. Note that there are two negative peaks (arrows) following the artefact, presumably representing active impulse propagation in afferent fibres with different conduction velocities, and that both were unchanged while the intracellular responses were potentiated. The first peak is nearly coincident with that of the coupling potential.

METHODS. In all experiments, standard surgical procedures⁷ were used; intracellular recordings (Rec) were obtained *in vivo* from the lateral dendrite (l. dend.) about 200 to 250 μ m from the M-cell soma, generally with a KCl (2.5 M) containing microelectrode of 7–15 M Ω resistance. In addition to eighth nerve stimulation (stim. 8th), the M-cell was also activated antidromically by an electrode on the spinal cord (stim. AD). All responses were measured, from prestimulation baseline to peak, after averaging sets of 15 or more traces. The eighth nerve test stimulus strength was such that the range of postsynaptic coupling potential amplitudes was ~15–25 mV and the chemical e.p.s.ps were about 5 mV. For subthreshold tetanization, conditioning and test stimulation strengths were comparable, whereas during supratherapeutic tetanization, the stimulus was sufficient to fire the M-cell at least once during each tetanus. In five experiments, p.s.ps were evoked by stimulation of the contralateral optic tectum, which projects to the ventral dendrite (v. dend.) of the M-cell. In another five experiments, a second electrode was located extracellularly, in the eighth nerve tract, about 400 μ m lateral to the M-cell soma and within 100 μ m of the dendrite.

levels of 43% and 57%, respectively. A similar study with the NMDA channel-blocker ketamine^{22–24} showed that at 50 μ M it blocked the potentiations of the coupling potential by 84% on average, and of the chemical e.p.s.p. by 86% ($n = 10$). At least two reasons could account for this incomplete block. First, in our *in vivo* preparation, the synapses are 1.5 mm below the surface of the brain, so the drug concentration at these sites might be lower than that in the superfusate. We did not try

using a higher concentration as any additional blocking effect might have been nonspecific²⁵. Second, the ketamine block is voltage- and use-dependent^{23,24}, and the strong supratherapeutic tetanization may have activated channels that were not yet blocked, thereby allowing some calcium influx and weak enhancements. Indeed, ketamine totally prevented the slower developing potentiations produced by subthreshold tetanization (Fig. 4*b, c*), a condition in which the test and conditioning

FIG. 2 Comparison of the long-term potentiations produced by subthreshold (*a, b*) and supratherapeutic (*c, d*) tetanizations of the eighth nerve. *a* and *b*, Time courses of changes in the coupling potential (*a*) and chemical e.p.s.p. (*b*) amplitudes ($n = 6$) following subthreshold tetanization (filled squares). The coupling potential increased only slightly compared with the control (open squares), while the e.p.s.p. showed a gradual increase which was significant by 15 min ($P < 0.05$). *c* and *d*, After supratherapeutic tetanization, however, both the coupling potential (*c*) and the e.p.s.p. (*d*) exhibited rapid and persistent increases ($n = 6$). For this and all subsequent figures, symbols represent the mean $\pm$ s.e.m.; for each experiment, the average response amplitude in the control period before tetanization was used as 100%. The period of tetanization is indicated by the shaded box on the abscissa before time zero.



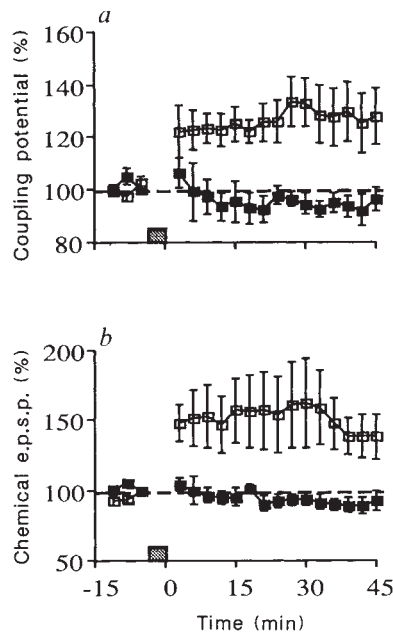


FIG. 3 Evidence for the involvement of calcium in the potentiations induced by suprathreshold tetanizations. *a* and *b*, Comparison of changes in the coupling potential (*a*) and chemical e.p.s.p. (*b*) obtained while recording with a microelectrode containing either (■) a calcium chelator, BAPTA, (10 mM in 2.5 M KCl; $n=7$), or (□) KCl only as control ($n=5$). The compound was injected iontophoretically into the M-cell with 5 nA negative current throughout the experiment, injections being interrupted only for the recordings. In the control, the same iontophoretic procedure was used but the electrode did not contain BAPTA.

stimuli were of comparable strengths. Thus, two selective NMDA antagonists that have different sites of action both blocked the induction of LTP. Finally, as NMDA channels are voltage-dependent as a result of magnesium blockage^{18,19}, the induction of LTP should have a similar dependency. By way of

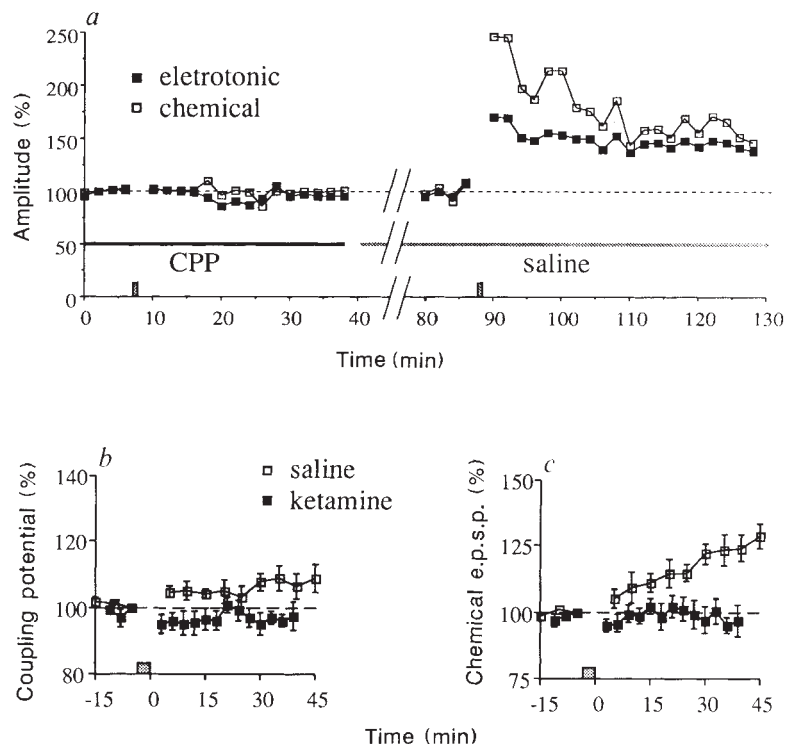
confirmation, we have observed that the induction of LTP is blocked by pairing a presynaptic tetanization, which normally produces potentiation, with a postsynaptic inhibition that shunts the synaptic depolarization (data not shown).

A subthreshold tetanization that activated fewer synapses without firing the postsynaptic cell might be expected not to produce LTP. But in the M-cell system the NMDA-mediated component contributes to the fast e.p.s.ps evoked at resting potential⁷, as has also been shown for single synapses on cerebellar granule cells²⁶. The depolarization that unblocks the NMDA channels^{18,19} might be produced by this e.p.s.p. and possibly the coupling potential as well. That is, subthreshold tetanization may activate some NMDA receptors, resulting in a slow-developing potentiation, and the larger effects produced by suprathreshold trains may not depend on the occasional M-cell spike initiation *per se*, but rather on the release of more transmitter.

A common mechanism could possibly account for the potentiation of each response. An increased presynaptic spike amplitude is unlikely, because after subthreshold tetanization the enhanced chemical e.p.s.p. was often not accompanied by a change in the coupling potential; moreover, as mentioned before, the extracellularly measured presynaptic spike never increased. These data and the indirect evidence that the M-cell input resistance remained constant throughout indicate that the potentiated coupling is due to an increased gap junctional conductance. The enhanced chemical e.p.s.p. could be related to an increase in transmitter release and/or postsynaptic responsiveness, but an increased duration of the presynaptic spike⁶ is probably not involved because the half-width of the coupling potential never increased after tetanization. Nevertheless, there is a common step in these two effects, because their inductions involve the activation of NMDA receptors and elevated intracellular calcium concentration. The observation that subthreshold tetanization produced a more significant potentiation in the chemical e.p.s.p. than in the coupling potential may reflect a subsequent divergence of the intracellular metabolic pathways or different sensitivities to a common intermediary.

An increase in intracellular calcium concentration generally decreases gap junctional conductance²⁷. As the potentiation

FIG. 4 Effects of NMDA-receptor antagonists on long-term potentiation of the eighth nerve responses. *a*, CPP (100 μ M in saline) prevented both potentiations after 10 trains of three suprathreshold stimuli at 500 Hz separated by 2-s intervals; after a 40-min washout period in saline, the same conditioning protocol produced immediate increases in both the coupling and the chemical e.p.s.p., which gradually decayed to elevated steady state levels. *b* and *c*, In the presence of ketamine (50 μ M in saline), the coupling potential decreased slightly after subthreshold tetanization (*b*), and the gradual increase in the e.p.s.p. was totally blocked (*c*; $n=6$ for both). Control curves correspond to the tetanization data in Fig. 2a and b, added here for comparison.



lasted much longer than the time period of tetanization, elevated intracellular calcium concentration probably does not itself have a role in modulating the conductance, but is more likely to be an intermediate step in the process.

These activity-dependent modifications of electrotonic coupling, and the fact that they occur at a sensory input to a command neuron, suggest that LTP is not unique to structures implicated in 'cognitive' functions. In the present case, they may produce sensitization of a primitive but vital escape reflex²⁸. □

Received 20 August; accepted 4 October 1990.

1. Bliss, T. V. P. & Lomo, T. *J. Physiol., Lond.* **232**, 311–356 (1973).
2. Lomo, T. *Acta Physiol. Scand.* **68**, (suppl. 277) 128 (1966).
3. Brown, T. H., Chapman, P. F., Kairiss, E. W. & Keenan, C. L. *Science* **242**, 724–728 (1988).
4. Gustafsson, B. & Wigstrom, H. *Trends Neurosci.* **11**, 156–162 (1988).
5. Dudek, F. E., Andrew, R. D., MacVicar, B. A., Snow, R. W. & Taylor, C. P. in *Basic Mechanisms of Neuronal Hyperexcitability* (eds Jasper, H. H. & Van Gelder, N. M.) 31–73 (Liss, New York, 1983).
6. Lin, J. W. & Faber, D. S. *J. Neurosci.* **8**, 1302–1325 (1988).
7. Wolszon, L. R. & Faber, D. S. *Soc. Neurosci. Abstr.* **14**, 939 (1988).
8. Fay, R. R. & Olsho, L. W. *Comp. Biochem. Physiol.* **62A**, 377–386 (1979).
9. Faber, D. S. & Korn, H. *Neurobiology of the Mauthner Cell* (eds Faber, D. S. & Korn, H.) 47–132 (Raven, New York, 1978).
10. Zottoli, S. J., Hordes, A. R. & Faber, D. S. *Brain Res.* **401**, 113–121 (1987).
11. Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. & F. Schottler, *Nature* **305**, 719–721 (1983).
12. Malenka, R. C., Kauer, J. A., Perkel, D. J. & Nicoll, R. A. *Trends Neurosci.* **12**, 444–450 (1989).
13. Tsien, R. Y. *Biochemistry* **19**, 2396–2404 (1980).
14. Williams, S. & Johnson, D. *Neuron* **3**, 583–588 (1989).
15. Collingridge, G. L., Kehi, S. J. & McLennan, H. *J. Physiol., Lond.* **334**, 33–46 (1983).
16. Collingridge, G. L. & Bliss, T. V. B. *Trends Neurosci.* **10**, 288–293 (1987).
17. Zalutsky, R. A. & Nicoll, R. A. *Science* **248**, 1619–1624 (1990).
18. Ascher, P. & Nowak, L. *J. Physiol., Lond.* **399**, 207–266 (1988).
19. Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* **394**, 501–528 (1987).
20. Davies, J. et al. *Brain Res.* **382**, 169–173 (1986).
21. Harris, E. W., Ganong, A. H., Monaghan, D. T., Watkins, J. C. & Cotman, C. W. *Brain Res.* **382**, 174–177 (1986).
22. Anis, N. A., Berry, S. C., Burton, N. R. & Lodge, D. *Br. J. Pharmac.* **79**, 565–575 (1983).
23. Honey, C. R., Miljkovic, Z. & MacDonald, J. F. *Neurosci. Lett.* **61**, 135–139 (1985).
24. MacDonald, J. F., Miljkovic, Z. M. & Pennefather, P. *J. Neurophysiol.* **58**, 251–266 (1987).
25. Benoit, E., Carrutu, M. R., Dubois, J. M. & Mitolo-Chieppa, D. M. *Br. J. Pharmac.* **87**, 281–297 (1986).
26. D'Angelo, E., Rossi, P. & Garthwaite, J. *Nature* **346**, 467–470 (1990).
27. Loewenstein, W. R. *Physiol. Rev.* **61**, 829–913 (1981).
28. Zottoli, S. J. *J. exp. Biol.* **66**, 243–254 (1979).

ACKNOWLEDGEMENTS. We thank J. Fetcho, M. J. Titmus and L. R. Wolszon for their help, and P. Ascher for comments on the manuscript. Supported in part by a grant from the NIH to D.S.F.

Neuropeptide modulation of single calcium and potassium channels detected with a new patch clamp configuration

Edwin S. Levitan^{*†} & Richard H. Kramer[‡]

^{*} Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

[‡] Howard Hughes Medical Institute and Center for Neurobiology and Behavior, Columbia University, College of Physicians and Surgeons, 722 West 168th Street, New York, New York 10032, USA

CALCIUM channel activity is crucial for secretion and synaptic transmission, but it has been difficult to study Ca^{2+} channel modulation because survival and regulation of some of these channels require cytoplasmic constituents that are lost with the formation of cell-free patches. Here we report a new patch clamp configuration in which activity and regulation of channels are maintained after removal from cells. A pipette containing the pore-forming agent nystatin is sealed onto a cell and withdrawn to form an enclosed vesicle. The resulting perforated vesicle, formed from pituitary tumour cells, contains Ca^{2+} and K^{+} channels. Ca^{2+} -activated K^{+} channels in the vesicle are activated by cyclic AMP analogues, and by a neuropeptide (thyrotropin-releasing hormone) that stimulates phosphatidylinositol turnover

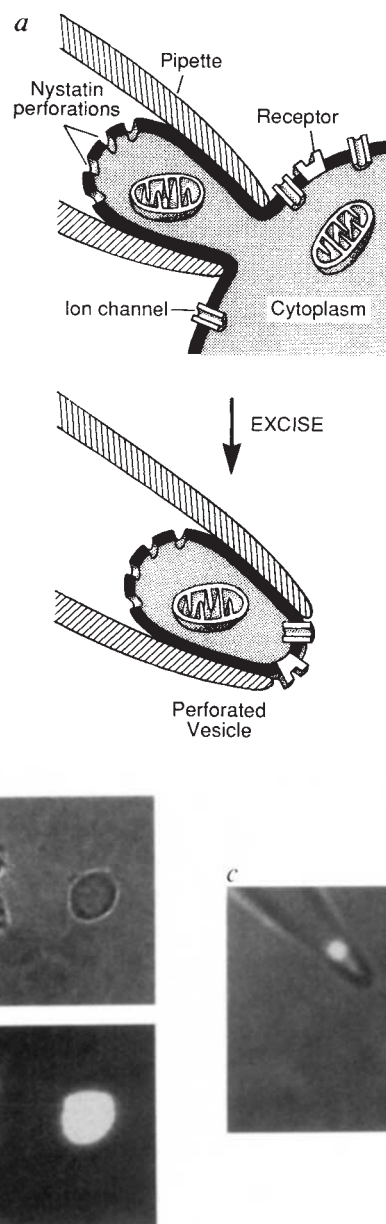
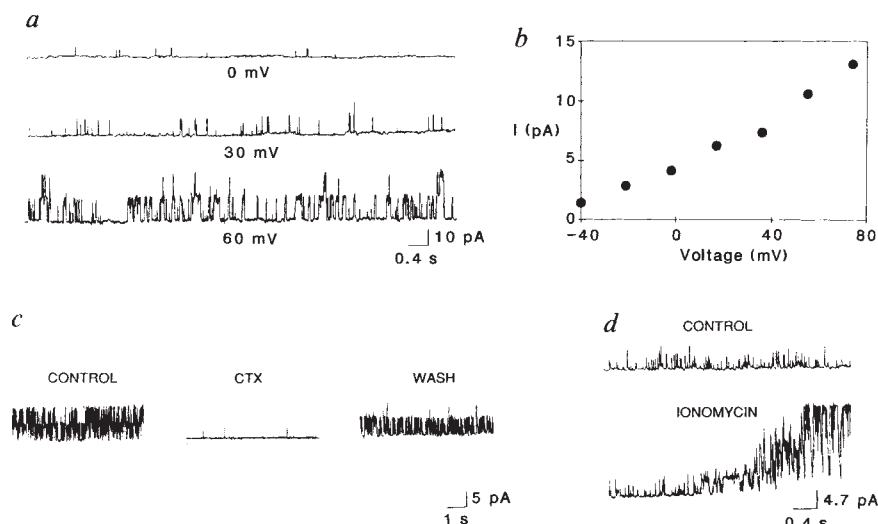


FIG. 1 The perforated vesicle. *a*, A cell-attached patch is formed with a pipette containing nystatin. After nystatin enters the membrane (top), the pipette is withdrawn to form a perforated vesicle (bottom). The membrane facing the bath contains receptors and ion channels with an outside-out orientation. *b*, Bright field (top) and fluorescent (bottom) images of GH3 cells loaded with 6-carboxyfluorescein (bar, 20 μm). *c*, Fluorescent image of a perforated vesicle in the tip of the pipette (bar, 10 μm). Dim white light was used to illuminate the pipette. While staining by 6-carboxyfluorescein was uniform, rhodamine 123 staining was punctate in both cells and vesicles. **METHODS.** GH3 or GH4C1 cells were grown in Ham's F10 medium supplemented with 15% horse serum, 2.5% fetal bovine serum, 2 mM glutamine, 50 U ml⁻¹ penicillin, and 50 μg ml⁻¹ streptomycin on Corning 35-mm tissue-culture dishes. During experiments the medium was replaced with 150 mM NaCl, 5.4 mM KCl, 2 mM CaCl₂, 0.8 mM MgCl₂, 20 mM glucose, 10 mM sodium HEPES, pH 7.5. The tip of the patch pipette was filled with solution I (150 mM KCl, 5 mM MgCl₂, 10 mM potassium HEPES, pH 7.1) or solution II (120 mM K₂SO₄, 16 mM KCl, 5 mM MgSO₄, 10 mM sodium HEPES, pH 7.1). The remainder of the pipette was then back-filled with the same solution II sonicated with 100 μg ml⁻¹ nystatin (stock solution 25 mg ml⁻¹ in dimethylsulphoxide). After formation of a gigaohm seal, progressive insertion of nystatin into the patch was monitored by measuring series resistance with a List EPC-7 patch clamp. When the series resistance became <50 Mohm, the clamp was briefly switched to current-clamp mode while the pipette was withdrawn to excise a vesicle. GH cells were loaded with 6-carboxyfluorescein by incubating cells with the nonfluorescent ester 6-carboxyfluorescein diacetate at 37 °C for 30 min.

[†] Present address: Department of Pharmacology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA.

FIG. 2 K[Ca] channels in perforated vesicles. *a*, Channel opening (upward deflections) increases with depolarization. *b*, Single-channel current versus voltage plot. Single-channel conductance was ~ 110 pS with solution I (data shown) and ~ 150 pS with solution II. The extrapolated reversal potential was about -65 mV, close to the expected value for E_K (-84 mV), consistent with a K^+ -selective channel. *c*, Effect of 50 nM charybdotoxin on channel activity. Voltage, $+10$ mV. *d*, Ionomycin (4 μ M) activates channels. Activation reversed with washing in 4 out of 5 cases. Voltage, $+10$ mV.



and inositol trisphosphate-gated Ca^{2+} release from intracellular organelles. Thus, the perforated vesicle retains signal transduction systems necessary for ion channel modulation. Functional dihydropyridine-sensitive Ca^{2+} channels (L-type) are maintained in the vesicle, and their gating is inhibited by thyrotropin-releasing hormone. Hence, this new patch clamp configuration has allowed a direct detection of the single-channel basis of transmitter-induced inhibition of L-type Ca^{2+} channels. The modulation of Ca^{2+} -channel gating may be an important mechanism for regulating hormone secretion from pituitary cells.

The procedure for obtaining the perforated vesicle configuration is shown in Fig. 1*a*. First, a cell-attached patch¹ is established on a cell with a pipette containing nystatin. Nystatin enters into the membrane and forms pores, which allow the recording of whole-cell electrical activity while maintaining cytoplasmic factors crucial for certain neurotransmitter responses^{2,3}. Withdrawing the pipette often results in the formation of an enclosed vesicle. The hemisphere of the vesicle in contact with nystatin inside the pipette remains perforated and provides a low-resistance pathway for measuring single-channel activity in the intact hemisphere in contact with the bath. To test whether perforated vesicles retain cytoplasm, GH3 or GH4C1 pituitary tumour cells were loaded with fluorescent dyes (Fig. 1*b*) and vesicles were excised. A water-soluble dye, 6-carboxyfluorescein (Fig. 1*c*), is clearly visible in the perforated vesicle, as is rhodamine 123, a fluorescent dye sequestered in mitochondria⁴. Thus, the vesicle retains water-soluble cytoplasmic elements and organelles such as mitochondria.

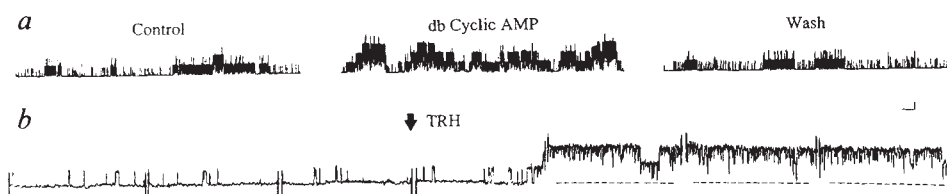
The vesicles contain ion channels with an outside-out orientation. One channel type activates on depolarization and has a large unitary conductance (Fig. 2*a, b*). This channel is blocked by addition of 4 – 8 mM tetraethylammonium ($N=4$) or 50 – 400 nM charybdotoxin ($N=4$), a peptide that blocks potassium

channels from the extracellular side of the membrane (Fig. 2*c*)⁵. The channels are also activated by elevating the internal Ca^{2+} concentration with the Ca^{2+} ionophore ionomycin ($N=5$) (Fig. 2*d*). Hence we identify this channel as the large conductance Ca^{2+} -activated K^+ (K[Ca]) channel^{6–8}.

K[Ca] channels are modulated by cAMP-dependent protein phosphorylation in many cells^{9–11}. The K[Ca] channels in the vesicle are modulated by application of membrane-permeant cAMP analogues; both dibutyryl cAMP ($N=2$ of 3) (Fig. 3*a*) and 8-Br-cAMP ($N=2$ of 4) reversibly stimulated K[Ca] channel activity. Direct application of cAMP on inside-out patches excised from GH3 cells failed to affect the activity of the K[Ca] channel ($N=4$). Hence we suggest that the perforated vesicle contains ATP and cAMP-dependent protein kinase, which mediates the effect of the cAMP analogues, as well as phosphoprotein phosphatase(s), which probably reverses the effect of the kinase.

Channels in the perforated vesicle can also be modulated by neurotransmitter-operated intracellular signalling systems that require functional organelles. In pituitary cells, the neuropeptide thyrotropin-releasing hormone (TRH) stimulates phosphatidylinositol hydrolysis, inositol trisphosphate-stimulated Ca^{2+} release from endoplasmic reticulum, and secretion of peptides^{12,13}. The increase in Ca^{2+} concentration also activates K[Ca] current^{7,14,15}. Application of TRH to perforated vesicles leads to a dramatic activation of K[Ca] channels ($N=5$). In Fig. 3*b*, before TRH application, no more than one channel is open throughout, and the channel openings are brief (<150 ms). After TRH, two channels are open almost continuously. In four of the five vesicles, K[Ca] activation was reversible, and in two of three attempts, responses were obtained with a second application of TRH. A similar activation of K[Ca] channels has been reported with cell-attached patches, but TRH has no effect on K[Ca] channels in conventional outside-out patches, in which

FIG. 3 Modulation of K[Ca] channels by activators of signal-transduction cascades. *a*, Activity of K[Ca] channels at $+30$ mV before, 1 min after bath application of 5 mM dibutyryl cAMP (db cAMP) and 2 min after washout of the analogue. Horizontal bar, 2 s; vertical bar, 7.5 pA. *b*, Sequential recordings of K[Ca] channel currents evoked by 2-s depolarizations (-50 mV to $+40$ mV) every 6 s before and after bath application of 1 μ M TRH. Horizontal bar, 200 ms; vertical bar, 6 pA.



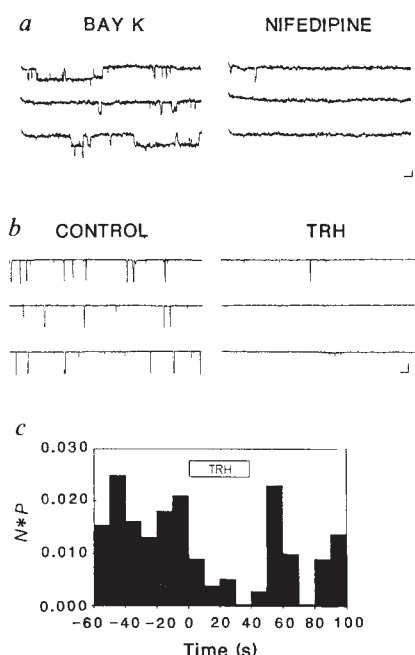


FIG. 4 Regulation of L-type Ca^{2+} channels by dihydropyridines and TRH. *a*, Current recordings from a vesicle with 2 μ M BayK 8644 and 1 μ M nifedipine. Inward current through L-channels (downward deflections) were evoked by depolarization from -40 mV to -20 mV every 2 s. Horizontal bar, 10 ms; vertical bar, 0.5 pA. *b*, Recordings of Ca^{2+} channel activity in the continued presence of 2 μ M BayK 8644 before (left) and 21 s after addition of 100 nM TRH (right) from another vesicle. Voltage, -25 mV. Horizontal bar, 100 ms; vertical bar, 0.55 pA. *c*, Time course of modulation of Ca^{2+} channels by TRH. Note that channel activity ($N \cdot P$, where N is the number of active channels and P is the probability of a channel being open) decreases with TRH application (horizontal bar) and recovers with washing. Data from same vesicle as in *b*.

METHODS. After excision of a perforated vesicle, the bath solution was replaced with 100 mM BaCl_2 , 20 mM TEA Cl, 10 mM caesium HEPES, pH 7.5. The dihydropyridine agonist was included to accentuate L-type channel activity. Whole cell recordings indicate that this agent does not affect the inhibition produced by TRH. GH cells contain T-type and L-type, but apparently no N-type, Ca^{2+} channels. TRH selectively inhibits L-type Ca^{2+} current¹⁸. Dihydropyridines in ethanol or dimethylsulphoxide were diluted into aqueous solutions; neither ethanol nor dimethylsulphoxide (0.02%) altered channel activity.

crucial cytoplasmic components have been apparently lost⁷. Hence, we conclude that all of the components of the signal transduction cascade, including the TRH receptor, G protein, phospholipase C, and intracellular inositol trisphosphate-releasable Ca^{2+} stores, are retained in the perforated vesicle.

Ca^{2+} channels are also found in the vesicle, and they too are modulated by TRH. L-type Ca^{2+} -channel activity from GH cells usually runs down in cell-free patches, unless artificially maintained in a phosphorylated state by addition of ATP and protein kinase, and inclusion of exogenous Ca^{2+} chelators¹⁶. In contrast, perforated vesicles display Ca^{2+} -channel activity that persists for at least 15 minutes, despite the absence of ATP or Ca^{2+} chelators in the patch pipette. These channels (~ 23 pS using Ba^{2+} as a charge carrier) are affected by dihydropyridines. The agonist BayK 8644 increases the duration and frequency of channel opening ($N = 12$), and the antagonists nimodipine and nifedipine prevent opening ($N = 7$) (Fig. 4*a*), characteristics of L-type Ca^{2+} channels¹⁷. We have recently found that TRH suppresses L-type Ca^{2+} current measured with the whole-cell perforated patch technique, but this regulation washes out with conventional whole-cell recording¹⁸. Likewise, L-channels in the vesicle are dramatically inhibited by TRH ($N = 5$ of 8). TRH decreases the frequency of channel opening without affecting the single-channel conductance (Fig. 4*b*). On removal of TRH, channel activity recovers (Fig. 4*c*). The decrease in frequency of opening could be due to a reduction in either the number of active channels in the patch, or a decrease in the rate of channel opening, or both. In two cases where overlapping openings were rare, we found a decrease in mean open time (2.9 ms to 1.4 ms for the vesicle shown). This modulation of L-type Ca^{2+} channels

is reminiscent of the inhibition of N-type Ca^{2+} channels in sympathetic neurons¹⁹. It will therefore be interesting to test whether inhibition of L-type channels is caused by a shift in the voltage dependence of gating, as has been proposed for inhibition of N-type Ca^{2+} channels²⁰.

The perforated vesicle affords advantages over previously described patch clamp configurations. With cell-attached patches the intracellular voltage is not controlled, and it is difficult to change the composition of the solution directly bathing the patch. With cell-free patches it is necessary to replace cytoplasm with an artificial solution that alters or disrupts ion channel function and regulation. In contrast, the perforated vesicle permits application of transmitters, toxins and drugs on channels that are unstable in outside-out patches. The perforated vesicle also allows the inspection of a small sample of a cell for the local presence of signal transduction cascades (that is, the volume of a vesicle is typically $<0.5\%$ of that of a GH cell). Thus, the observation that K^+ and Ca^{2+} channels are modulated by TRH in the vesicle indicates that the receptor, intracellular messengers, organelles and ion channels are all in close proximity.

The activity of L-type Ca^{2+} channels is important for secretion of peptide hormones from pituitary cells²¹⁻²³. The perforated vesicle has revealed for the first time inhibition of single L-type Ca^{2+} channels by a neuromodulator (TRH). TRH triggers release of internal Ca^{2+} and activation of protein kinase C (refs 12, 13). A comparison of the effects of selective activation of each of these messengers with the effects of TRH will help elucidate the mechanism of Ca^{2+} -channel inhibition. □

Received 17 July; accepted 14 September 1990.

- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch.* **391**, 85-100 (1981).
- Horn, R. & Marty, A. *J. gen. Physiol.* **92**, 145-159 (1988).
- Korn, S. J. & Horn, R. *J. gen. Physiol.* **5**, 789-812 (1989).
- Johnson, L. V., Walsh, M. L. & Chen, L. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 990-994 (1980).
- Miller, C., Moczydlowski, E., Latorre, R. & Phillips, M. *Nature* **313**, 316-318 (1985).
- Marty, A. *Nature* **291**, 497-500 (1981).
- Dubinsky, J. M. & Oxford, G. S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4282-4286 (1985).
- Lang, D. G. & Ritchie, A. K. *Pflügers Arch.* **410**, 614-622 (1987).
- Ewald, D. A., Williams, A. & Levitan, I. B. *Nature* **315**, 503-506 (1985).
- Lechleiter, J. D., Dartt, D. A. & Brehm, P. *Neuron* **1**, 227-235 (1988).
- Sikdar, S. K., McIntosh, R. P. & Mason, W. T. *Brain Res.* **496**, 113-123 (1989).
- Drummond, A. H. *J. exp. Biol.* **124**, 337-358 (1986).
- Gershengorn, M. C. *A. Rev. Physiol.* **48**, 515-526 (1986).
- Ritchie, A. K. *J. Physiol.* **385**, 611-625 (1987).

- Mollard, P., Vacher, P., Dufy, B., Winiger, B. P. & Schlegel, W. *J. biol. Chem.* **263**, 19570-19576 (1988).
- Armstrong, D. & Eckert, R. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2518-2522 (1987).
- Bean, B. P. *A. Rev. Physiol.* **51**, 367-384 (1989).
- Kramer, R. H., Levitan, E. S. & Kaczmarek, L. K. *Soc. Neurosci. Abstr.* **15**, 178 (1989).
- Lipscombe, D., Kongsamut, S. & Tsien, R. W. *Nature* **340**, 639-642 (1989).
- Bean, B. P. *Nature* **340**, 153-156 (1989).
- Albert, P. R. & Tashjian, A. H. *J. biol. Chem.* **259**, 15350-15363 (1984).
- Cronin, M. J. *et al. Am. J. Physiol.* **249**, E326-329 (1985).
- Enyeart, J. J., Aizawa, T. & Hinkley, P. M. *Am. J. Physiol.* **248**, C510-519 (1985).

ACKNOWLEDGEMENTS. We thank L. K. Kaczmarek and S. Siegelbaum for their support and comments on the manuscript; P. Darnies for GH4C1 cells; F. Sigworth for use of his computer analysis system; H. Blumenfeld for assistance with imaging; A. Scriabine (Miles Labs) for nimodipine and Bay K; and C. Miller and E. Moczydlowski for charybdotoxin. This research was supported by the NIH (E.S.L., L.K.K.), the Grass Foundation (E.S.L.) and by the Howard Hughes Medical Institute (R.H.K.).

Laser inactivation of fasciclin I disrupts axon adhesion of grasshopper pioneer neurons

Daniel G. Jay* & Haig Keshishian†

* Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138, USA

† Department of Biology, Yale University, New Haven, Connecticut 06511, USA

A MOLECULAR mechanism for selective axonal adhesion is a central question of neural development. Cell adhesion molecules have been identified¹, but it has been difficult to ascribe functions for these proteins *in vivo*. Here we show that the neuronal membrane glycoprotein fasciclin I has a role in the adhesion of sister axons during the development of the grasshopper limb bud. To do this we used a new technique, chromophore-assisted laser inactivation (CALI), which causes the precisely timed thermal denaturation of specific proteins by laser light targeted through a dye-labelled antibody, without any other observable damage to living cells². This can be achieved by relaxation of the laser-excited dye which releases heat to denature the bound protein; the rapid dissipation of heat with distance insulates unbound proteins from damage. CALI is a molecular analogue of cellular laser ablation and provides an unprecedented level of spatial and temporal resolution. Using dye-labelled antibodies that recognize fasciclin I, CALI disrupts fasciculation of the pioneer neurons without affecting their growth or guidance.

Neuron-specific membrane proteins in the grasshopper and *Drosophila*, called fasciclins, are expressed in a stage-specific

FIG. 1 General method for CALI with grasshopper eggs. *a*, Injection of dye-labelled reagents into grasshopper embryos. Ant., anterior. *b*, Laser irradiation of injected grasshopper embryos. *c*, T1 neurons (arrowhead) at the time of reagent injection are newly emerged from the epidermis and have not begun axonogenesis. The embryo shown was fixed and immunostained at the time that its sibling embryos were injected. This was to facilitate visualization; for CALI, neurons were staged using Nomarski optics alone. Scale bar, 20 μ m. *d*, T1 axon trajectory following 44 h incubation after injection of the embryo with anti-HRP without dye and subjected to laser irradiation. It was indistinguishable from untreated embryos^{3,5}. The extent of growth is demarcated by the numbers from 1 to 5, proximally to distally from the cell body in order to quantitate axon extension (see Table 1). Scale bar, 50 μ m.

METHODS. Antibodies and other reagents were labelled with Malachite green isothiocyanate according to ref. 2. A solution of Malachite green isothiocyanate (Molecular Probes) was freshly prepared by dissolving the dry reagent with dimethylsulphoxide at a concentration of 10 mg ml⁻¹. The dye was added to a solution of 400 μ g of test reagent in 500 μ l 0.5 M NaHCO₃, pH 9.5, in four 10- μ l aliquots once every 10 min with continuous rocking. The mixture was incubated for an additional 30 min. The labelled reagent was separated from free label by gel filtration with a prepacked PD-10 column (Pharmacia) eluted with Tris-buffered saline pH 7.4. The final concentration of the antibody was 200 μ g ml⁻¹. The ratio of labelling was obtained by measuring the optical density of the labelled solution at 620 nm (molar absorptivity, 150,000 M⁻¹ cm⁻¹) to determine the concentration of dye and dividing that value by the reagent concentration. The following reagents were labelled (dye:ligand ratio in parentheses): BSA (8), hexaalanine (1), goat anti-HRP IgG fraction (4) (Cappel) and mouse monoclonal anti-fasciclin I in ascites fluid (3) (a gift from A. Harrelson and C. Goodman, University of California, Berkeley). Grasshopper eggs were staged at 31 $\pm$ 1% of development by imaging the newly born T1 neurons using Nomarski optics. Staging was confirmed by immunostaining with anti-HRP. Grasshopper eggs were microinjected with 0.5 μ l reagent in phosphate-buffered saline using a 30-micron-tip glass electrode by 3 $\times$ 50

TABLE 1 Effect of CALI on axonal growth

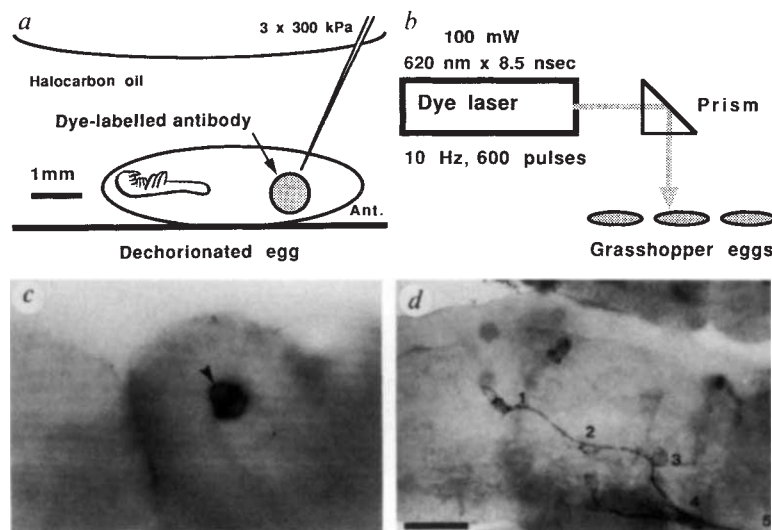
Treatment	Axon length*	Standard deviation	Number of cases
MG-BSA	4.0	0.89	5
MG-BSA + laser	4.41	1.08	17
anti-HRP	4.8	0.4	5
anti-HRP + laser	4.08	1.19	25
MG-anti-HRP	3.8	0.91	17
MG-anti-HRP + laser	4.37	1.07	59

* Axonal growth was tabulated by position along the axonal pathway marked 1–5 in Fig. 1*d*. The value 5 is given to axons that have reached the CNS.

The metathoracic limb buds examined were from sibling embryos from the same experiment. Similar results were obtained with the prothoracic and mesothoracic limb buds. T1 neurons from embryos are compared after injection with MG-BSA, anti-HRP, or MG-anti-HRP, with or without laser irradiation.

manner and localize to a subset of neurons during development^{3,4}. Fasciclin I is a cell adhesion molecule of relative molecular mass 70,000 that shows no homology with any known protein⁵. Although fasciclin I promotes cell adhesion when expressed on S2 *Drosophila* cells, showing homophilic binding⁶, the role for this protein *in vivo* remains to be demonstrated.

A simple system to investigate fasciclin I function is found in the T1 pioneers, the first neuron pair to differentiate in the grasshopper limb bud⁷. Fasciclin I is expressed on the surface of T1 neurons during their development³. These cells extend their axons along a stereotypic pathway toward the central nervous system (CNS), crossing the epithelium at a time when no other axons are present^{8,9}. The cell bodies, axons and growth



millisecond pulses of nitrogen from a Picospritzer II microinjector (General Valve Corp.) set at 300 kilopascals. We and others¹⁵ have shown by immunostaining that the antibodies reach the T1 neurons. The eggs were incubated for 6 to 12 h in halocarbon oil (700 W) in a humidified 31 $^{\circ}$ C incubator and then subjected to laser irradiation. The second harmonic frequency (532 nm) of a Quanta Ray DCR 3 Nd:YAG laser was used to drive a Quanta Ray PDL2 dye laser containing the fluorescent laser dye, DCM (4(dicyanomethylene)-2-methyl-6 (*p*-dimethylaminostyryl)-4H-pyran), to generate a 620-nm pulsed laser beam with a pulse duration of 8.5 nanoseconds at a frequency of 10 Hz. The beam was focused with an interjected convex lens and directed to a slide containing the eggs using a right angle prism. The spot diameter on the samples was 2 mm. The pulse energy was 10 mJoules. Samples were subjected to 1 min laser pulsing (600 pulses) for a total delivered energy of 6 joules.

cones of the sister neurons remain in close contact with each other. Aberrant behaviour of the Ti1 pioneers can be deduced from their axon trajectories.

Fasciclin I was inactivated by CALI using horseradish peroxidase (HRP), which fortuitously recognizes this protein^{10,11} or anti-fasciclin I monoclonal antibody conjugated with Malachite green. This dye absorbs light at 620 nm, a wavelength not significantly absorbed by cells. Grasshopper eggs were injected with either dye-labelled or unlabelled antibodies, dye-labelled bovine serum albumin (BSA), or dye-labelled hexaalanine (Fig. 1a) and incubated for 4–6 hours to label the neurons vitally. Subsets of these embryos were laser-irradiated (Fig. 1b). Embryos were then incubated for 24 h to allow for axon extension. The embryos were fixed and immunostained with anti-HRP¹⁰ to observe the axonal projections of the Ti1 pioneers. Figure 1c shows immunostained Ti1 neurons at the time of antibody injection, just before initial axonal outgrowth; Fig. 1d shows a limb bud from a control embryo that was immunostained 44 h after injection with unlabelled anti-HRP and laser irradiation. The sister axons of the Ti1 neurons appear as a single fascicle with a correctly guided trajectory.

Figure 2 shows the T3 limb buds from sibling embryos injected with Malachite green-labelled anti-HRP and subjected to either

no laser irradiation (Fig. 2a) or to 600 pulses of laser light (Fig. 2b). The Ti1 neurons in Fig. 2a are indistinguishable from those observed in untreated embryos. Laser irradiation resulted in defasciculation (Fig. 2b). Figure 2c and d illustrates a similar result with CALI using anti-fasciclin I. The separated axons rejoin after growing 50–100 micrometres, often near guideposts or at segment boundaries. There is a correlation of the position of defasciculation with the time of laser irradiation, suggesting that CALI inactivates fasciclin I on the growth cones, resulting in separation of the two axons.

The results of 1,267 limb buds are shown in Fig. 3a. Normally one in six pioneer neurons defasciculate in untreated embryos. This frequency is increased 2–3 fold by CALI using either

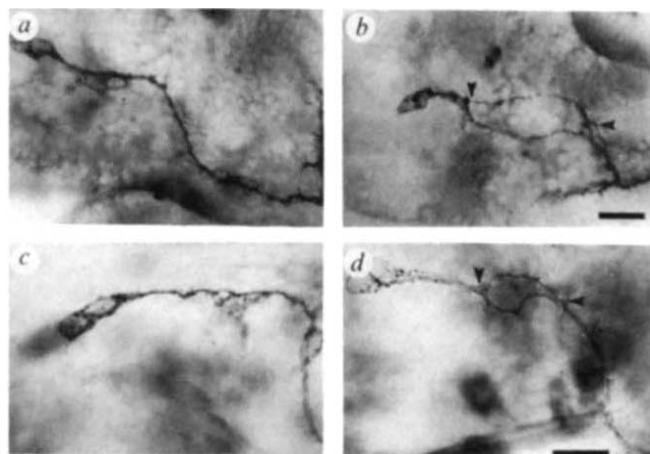


FIG. 2 CALI causes the defasciculation of the Ti1 sister axons. The arrows indicate the points of axons separating in CALI treated embryos (b and d); these axons remain as single fascicles in embryos injected with dye-labelled antibodies but not subjected to laser irradiation (a and c). Ti1 neurons are shown from the metathoracic limb buds of sibling embryos that have been injected as follows: a, with Malachite green-labelled anti-HRP without laser irradiation; b, with Malachite green-labelled anti-HRP and subjected to laser irradiation; c, with Malachite green-labelled anti-fasciclin I without laser irradiation; d, with Malachite green-labelled anti-fasciclin I subjected to laser irradiation. Scale bar in a, b, 30 μ m; in c, d, 40 μ m.

METHODS. After CALI, embryos were incubated at 31 °C for an additional 24–36 h. They were then fixed for 2 h in cold 4% paraformaldehyde dissolved in 100 mM NaH_2PO_4 , adjusted to pH 7.4. After fixation, embryos were washed with agitation for 1 h in cold buffered Triton-saline (0.3% Triton X-100 in 20 mM NaH_2PO_4 , 150 mM NaCl, pH 7.3). All subsequent solutions were prepared in this buffer and run at 4 °C with agitation unless indicated. Nonspecific staining was blocked with a 1-hour incubation with agitation in 1% BSA. Neurons were labelled by an overnight incubation with a rabbit derived anti-HRP IgG after CALI treatment, followed by a fluorescein isothiocyanate (FITC)-labelled goat anti-rabbit¹⁷. Neurons remain immunoreactive to anti-HRP after CALI treatment. As anti-HRP recognizes a glycosyl moiety¹⁶, denaturation should not disrupt immunoreactivity. Embryos were washed for a further hour at room temperature and mounted in 5% propyl gallate in glycerol, pH 8.0. Each grasshopper embryo was mounted, ventral side up, in mountant and gently flattened with the coverslip to hold the six limb buds outward. Embryos were stored for up to a year frozen at –80 °C with little loss of staining. The FITC-anti-HRP labelled neurons were imaged using digital optical microscopy as described¹⁷.

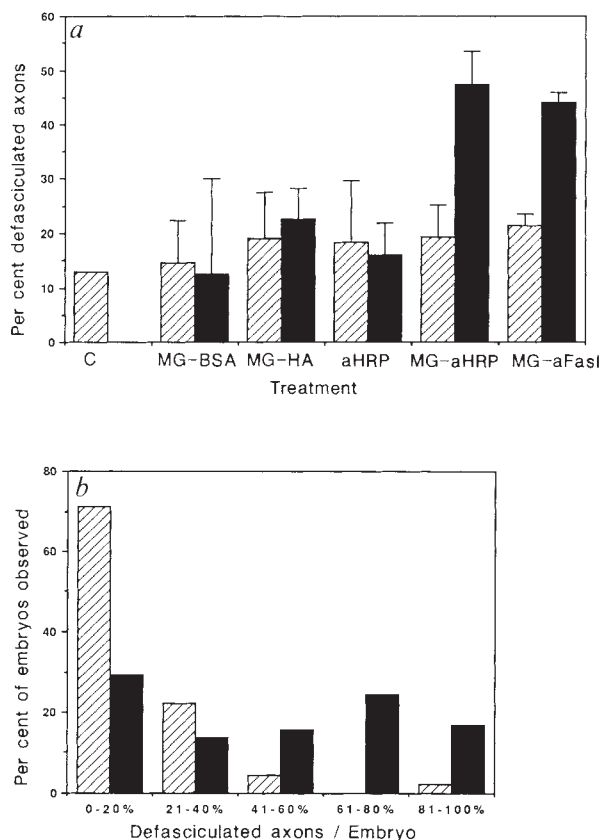


FIG. 3 The effect of CALI on the defasciculation of Ti1 neurons. a, CALI causes an increase in the rate of defasciculation, whereas control embryos are not significantly different from untreated embryos. The Ti1 neurons were imaged and separations of the sister axons that were greater than 25 microns in length were scored as a defasciculation event. Data are expressed as the percentage of limb buds that showed a defasciculation event for each experimental treatment. Each value is the average between experiments, except for Malachite green (MG)-BSA injected embryos, which are the average between pods, and the untreated embryos, which were derived from a single pod. Samples subjected to laser irradiation are shown in black bars. The total number of limb buds examined for each experimental class (–laser, +laser) are shown in parentheses: C, untreated embryos (–laser, 63); MG-BSA: embryos injected with dye-labelled BSA (–laser, 16; +laser, 46); MG-HA: embryos injected with dye-labelled hexaalanine (–laser, 48; +laser, 230); aHRP: embryos injected with anti-HRP without dye (–laser, 81; +laser, 230); MG-aHRP: embryos injected with dye-labelled anti-HRP (–laser=128, +laser=322); MG-aFasI: embryos injected with dye-labelled anti-fasciclin I (–laser, 94; +laser, 186). b, CALI increases the number of defasciculated axons observed per embryo. Data are expressed as the percentage of limb buds showing defasciculated axons for each embryo. The axons of all six limb buds were not always observable. The hatched bars represent embryos injected with anti-HRP antibody and subjected to laser irradiation ($n=45$). The black bars represent embryos injected with Malachite green-labelled anti-HRP and subjected to laser irradiation ($n=65$).

anti-fasciclin I or anti-HRP. By contrast, unirradiated embryos injected with dye-labelled antibodies, or embryos injected with unlabelled antibodies, regardless of laser irradiation, showed low defasciculation rates. No effect was seen with laser irradiation of embryos injected with labelled BSA, which presents the dye on a nonspecific protein that does not bind membranes. Similarly, there was no effect when labelled hexaalanine was used; this is a hydrophobic peptide that binds to membranes, but not specifically to proteins, thereby acting as a control for nonspecific damage to the membrane by CALI. Thus, it is insufficient to have the dye present on the cell surface to cause defasciculation; the dye must be linked to a specific protein. CALI-treated embryos also had high defasciculation rates on a per-individual basis. Figure 3b compares the percentage of defasciculated trajectories per animal in laser-treated anti-HRP (open bars) versus laser-treated Malachite green/anti-HRP-injected embryos (filled bars). Malachite green/anti-fasciclin I laser-treated embryos showed a similar distribution to that seen using Malachite green/anti-HRP.

Over 80% of injected embryos survive CALI. The treated limb buds show normal segmentation and growth, with the overall length of the T1 axons unaffected (Table 1). The growth cones also appear normal; even defasciculated neurons contact guidepost cells, cross leg segment boundaries and show typical filopodial spread¹². New neurons, such as the femoral chordotonal organ and subgenual organ¹³, are able to differentiate after CALI. CALI does not influence the differentiation of guidepost cells ($n = 202$ limb buds from a single experiment). As CALI causes defasciculation without affecting other phenomena, T1 elongation and guidance are not dependent on fasciculation.

Is fasciclin I solely responsible for the axon adhesion observed in these experiments? Elkins *et al.*¹⁴ have shown that *Drosophila* null mutations of fasciclin I have minor effects on the CNS, but that double mutants of fasciclin I and the Abelson tyrosine kinase gene show significant disruption in axonal organization. As CALI can indirectly inactivate a small complex *in vitro* (data not shown), it is possible that a protein closely bound to fasciclin I is also inactivated. Alternatively, regulatory mechanisms may compensate for the loss of fasciclin I in null mutations. This may not occur during the acute inactivation resulting from CALI.

Cell adhesion has been difficult to demonstrate *in situ* during neuronal development because of the paucity of effective inhibiting probes, and the difficulty of controlling their action over a specific time interval. We have employed a novel technique to demonstrate the role of a specific molecule in mediating axon adhesion by converting a binding reagent into an inhibitor. We suggest that CALI could be generally applied in the functional inactivation of other proteins. □

Interleukin-1 receptor antagonist reduces mortality from endotoxin shock

Kjell Ohlsson*, Peter Björk*, Magnus Bergenfeldt*, Robert Hageman† & Robert C. Thompson†

* Department of Surgical Pathophysiology, University of Lund, Malmö General Hospital, S21401, Malmö, Sweden

† Synergen, Inc., 1885 33rd Street, Boulder, Colorado 80301, USA

ABOUT five out of 1,000 patients admitted to hospital develop bacterial sepsis leading to shock¹, the mortality rate for which is high despite antibiotic therapy². The infection results in hypotension and poor tissue perfusion, and eventually leads to the failure of several organ systems. Bacterial endotoxin is thought to be the direct cause of shock in Gram-negative sepsis, because it can cause shock in animals³, and antibodies against endotoxin prevent Gram-negative shock in animals⁴ and in humans⁵⁻⁷. But, the symptoms of septic shock are the result of the actions of host cytokines induced by the endotoxin. The cytokine interleukin-1 has been implicated as an important mediator of septic shock because it can induce tachycardia and hypotension and act synergistically with tumour necrosis factor to cause tissue damage⁸ and death⁹. We now report that a specific interleukin-1 receptor antagonist reduces the lethality of endotoxin-induced shock in rabbits, indicating that interleukin-1 does indeed play an important part in endotoxin shock.

A recombinant human interleukin-1 (IL-1) receptor antagonist (IL-1ra)^{10,11} that blocks the effects of IL-1 *in vitro* has provided a tool for determining the role of IL-1 in animal models of septic shock. The hypotension and leukopenia that follow a single intravenous injection of recombinant human interleukin-1 β (IL-1 β , 15 $\mu\text{g kg}^{-1}$) into anaesthetized rabbits⁸ were blocked in a dose-dependent way by the injection of IL-1ra shortly before IL-1 (Fig. 1a, b). The leukocytosis that occurs at later times in conscious rabbits injected with IL-1 β was also blocked (Fig. 1c). These results indicate that IL-1ra should block the effects of endotoxin shock mediated by IL-1 in rabbits.

Because death is a common outcome of septic shock in humans, we tested the effects of IL-1ra in a model of endotoxin shock in rabbits, for which the mortality rate is similarly high. Intravenous injection of *Escherichia coli* endotoxin (0.5 mg kg⁻¹) into rabbits killed eight of ten animals within 48 hours (Fig. 2). Within the first few hours after the injection the animals' fur became ruffled and they became essentially immobile. The animals' breathing became strained and harsh noises could be heard issuing from their lungs. At autopsy, the lungs showed striking pathological changes. On gross examination they were heavy with a liver-like appearance. Light microscopy revealed that the fine alveolar architecture was disrupted, showing oedematous alveolar walls and a massive accumulation of protein and of red and white blood cells (Fig. 3a).

By contrast, nine of ten rabbits receiving a total of 100 mg kg⁻¹ IL-1ra and the endotoxin survived the observation period of 7 days and appeared to make a full recovery (for dosage, see Fig. 2). These animals moved about the cage, and their fur and breathing appeared to be normal, although the animals were less lively than untreated controls. Rabbits of another group treated in the same way were killed after 24 hours so that their lungs could be examined. On gross examination, small areas of surface bleeding but no gross hepatization was seen, and on microscopic examination, there was no evidence of the massive transudation and cellular infiltration seen in the group treated with endotoxin alone (Fig. 3b). The beneficial effects of IL-1ra are dose-dependent. A group of 10 rabbits treated with 30 mg kg⁻¹ IL-1ra had an intermediate level of mortality, and

Received 4 July; accepted 4 October 1990.

- Jessell, T. M. *Neuron* **1**, 3-13 (1988).
- Jay, D. G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5454-5458 (1988).
- Bastiani, M. J., Harrelson, A. L., Snow, P. M. & Goodman, C. S. *Cell* **48**, 745-755 (1987).
- Patel, N. H., Snow, P. M. & Goodman, C. S. *Cell* **48**, 975-988 (1988).
- Zinn, K., McAllister, L. & Goodman, C. S. *Cell* **53**, 577-587 (1988).
- Elkins, T., Hortsch, M., Bieber, A., Snow, P. & Goodman, C. S. *J. Cell Biol.* **110**, 1825-1832 (1990).
- Bate, C. M. *Nature* **260**, 54-55 (1976).
- Keshishian, H. & Bentley, D. *Dev. Biol.* **96**, 89-102 (1983).
- Condic, M. L. & Bentley, D. *J. Neurosci.* **9**, 2678-2686 (1989).
- Jan, L. Y. & Jan, Y. N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2700-2704 (1982).
- Snow, P. M., Patel, N. H., Harrelson, A. L. & Goodman, C. S. *J. Neurosci.* **7**, 4137-4144 (1987).
- Caudy, M. & Bentley, D. *J. Neurosci.* **6**, 1781-1795 (1986).
- Keshishian, H. & Bentley, D. *Dev. Biol.* **96**, 103-115 (1983).
- Elkins, T. *et al. Cell* **60**, 565-575 (1990).
- Petry, D., Buster, D., Donato, K. K. & Anderson, H. *Dev. Growth Differ.* **31**, 299-305 (1989).
- Katz, F., Moats, W. & Jan, Y. N. *EMBO J.* **7**, 3471-3477 (1988).
- Johansen, J., Halpern, M. E. & Keshishian, H. *J. Neurosci.* **9**, 4318-4332 (1989).

ACKNOWLEDGEMENTS. We thank D. Bentley, C. Goodman, W. Gilbert and R. Wyman for comments on the manuscript. A. Harrelson and C. Goodman for the anti-fasciclin I monoclonal, and the MIT Laser Research Center for use of facilities. Supported by the Lucille P. Markey Charitable Trust, the Whitaker Foundation, and Hoffmann La Roche (D.J.), and by the NIH and the March of Dimes Birth Defects Foundation (H.K.).

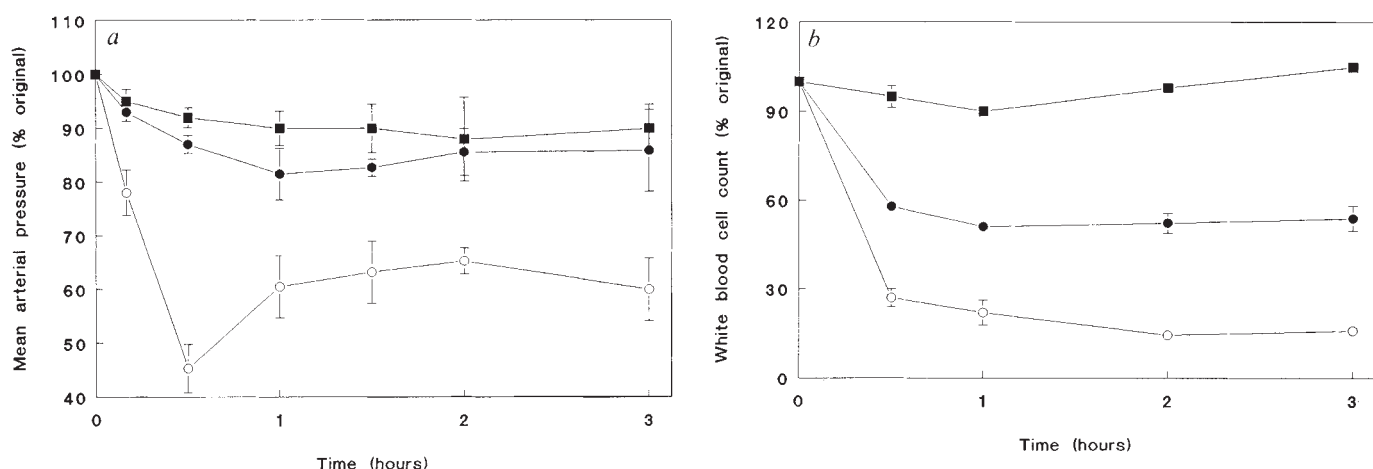
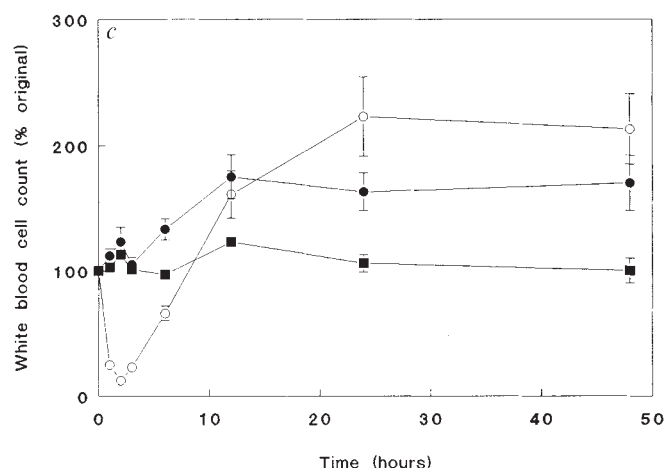


FIG. 1 Effect of IL-1ra on IL-1 β -induced changes in mean arterial pressure and white blood cell count in anaesthetized and conscious rabbits, which are as sensitive to endotoxin and IL-1 as humans⁸. Rabbits (blue chinchilla, 2.5 kg; BomMice, Bomholtgaard Breeding and Research Center Ltd, Bomholtvej 10, DK-8680, Ry, Denmark) were anaesthetized with a single injection of 4 mg kg⁻¹ xylazine and 10 mg kg⁻¹ ketamine. Catheters (PE 50) were placed in the left carotid artery and the superior caval vein to allow for the continuous recording of the arterial and central venous pressure. IL-1 β and IL-1ra (both prepared from *E. coli* and containing <0.5 U endotoxin per mg protein) were injected into the central venous catheter as a bolus over 1 min. During the observation period, blood was withdrawn from the carotid artery catheter for white blood cell count (originally close to 7×10^9 L⁻¹) and platelet counts, and directly into EDTA for plasma samples. Blood samples removed from the catheters were replaced with the same volume of saline (~5 ml over 3 h). The total volume of saline administered was 6 ml per kg body weight per hour. In one series of IL-1/IL-1ra experiments (c) the animals were allowed to wake up after the catheters were in place. The catheters were secured in the neck to allow repeated measurements of blood pressure and also sampling of blood. Rabbits received either 15 μ g kg⁻¹ IL-1 β (○), 15 μ g kg⁻¹ IL-1 β and 1 mg kg⁻¹ IL-1ra (●), or 15 μ g kg⁻¹ IL-1 β and 4 mg kg⁻¹ IL-1ra (■). Rabbits injected with saline or with 4 mg kg⁻¹ IL-1ra remained haemodynamically stable and showed no significant change in the number of white blood cells throughout the observation period (data not shown).

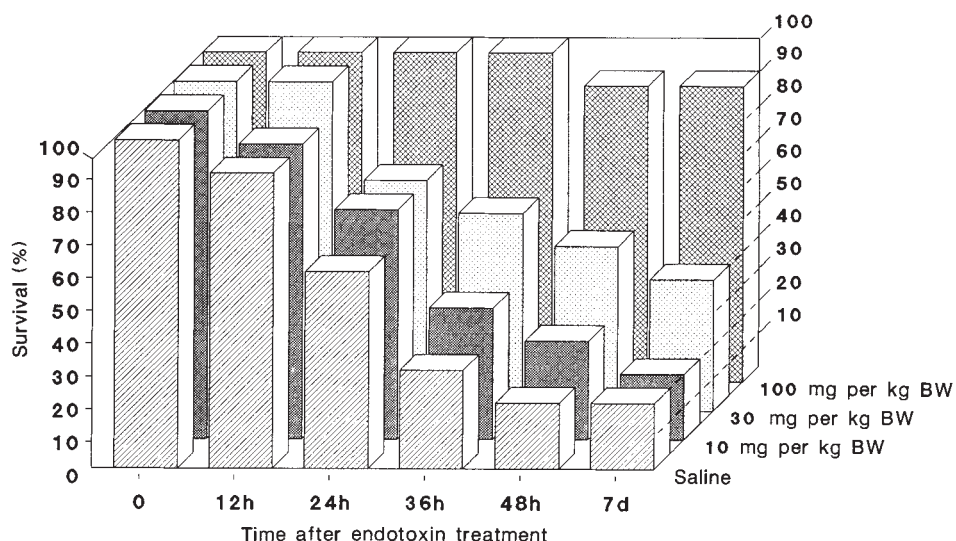


animals given 10 mg kg⁻¹ IL-1ra had the same level of mortality as untreated animals (Fig. 2).

Large doses of IL-1ra were required to block the action of IL-1 in the IL-1- and the endotoxin-induced diseases, despite the fact that IL-1ra and IL-1 are expected to have similar affinities for the IL-1 receptor on the basis of experiments with

mouse cells¹⁰. In part this high dose indicates that more than 50% of IL-1 binding needs to be blocked for 50% of the biological effects of IL-1 to be blocked. Similar effects occur with cells in culture¹². But the high doses needed are also a consequence of the rapid equilibration and clearance of IL-1ra after intravenous injection. In the group of animals receiving a total of 100 mg

FIG. 2 The effect of IL-1ra on the survival rate in endotoxin-induced shock in rabbits. IL-1ra was injected in equal doses just before the injection of endotoxin (*E. coli* 026, B26; Sigma) and every 2 h thereafter for 24 h. The rabbits were not anaesthetized but were under constant observation for 48 h and were then observed during the daytime for up to 7 days. The animals had free access to water and food. Injections and blood sampling were made through the ear veins. BW, body weight; d, days.



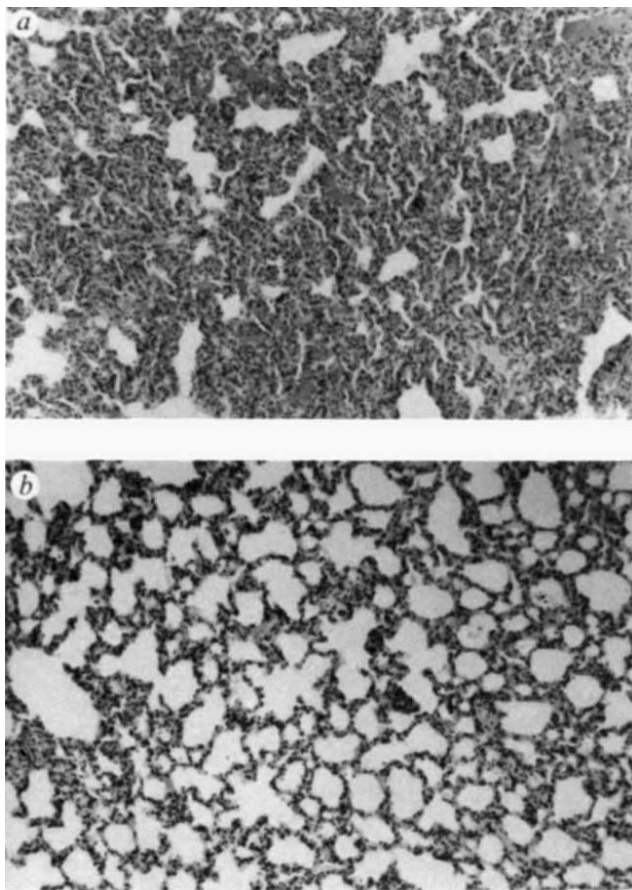


FIG. 3 Microscopic appearance (magnification, 160 $\times$) of a rabbit lung removed: a, immediately after death at 26 h from an animal receiving endotoxin and no IL-1ra; and b, 24 h after identical injection of endotoxin from an animal receiving 100 mg rhIL-1ra per kg body weight over the first 24 h. Rabbits were killed with pentobarbital. The tissue specimens were fixed in buffered (pH 7.4) 10% formalin, dehydrated and embedded in paraffin. The tissue samples were analysed with light microscopy with haematoxylin-eosin staining.

IL-1ra per kg body weight, the level of IL-1ra in plasma, as measured by a single radial immunodiffusion assay¹³, varied from between 150 $\mu\text{g ml}^{-1}$ 5 minutes after each 7.7-mg kg^{-1} dose, and 20 $\mu\text{g ml}^{-1}$ 2 hours later. These results imply that the quantity of IL-1ra needed to prevent mortality would be considerably reduced if the circulation of the protein could be prolonged.

To determine the time during which IL-1 acts a pathological agent in endotoxin shock we investigated the effects on mortality of delaying the treatment with IL-1ra. When the standard treatment with IL-1ra was delayed for 1 or 2 hours after the endotoxin injection, seven of the eight animals in each group survived the 7-day period whereas three of the four animals in a group not treated with IL-1ra died within 48 hours. These results indicate that IL-1 toxicity can be reversed at least 2 hours after injecting endotoxin. Because IL-1 seems to be a late-acting agent in endotoxin shock, IL-1ra could have therapeutic as well as prophylactic properties in septic shock. This is especially important clinically, where the disease, as measured by hypotension, can be in progress before intervention can be initiated. Experiments are in progress to determine the latest time after administration of endotoxin at which the disease can be prevented.

In light of earlier results implicating tumour necrosis factor as a mediator in endotoxin shock¹⁴⁻¹⁷, the demonstration that IL-1 is also an important mediator shows that the disease probably results from several cytokines acting with additive or syner-

gistic effects. This conclusion is in accordance with IL-1 greatly potentiating the shock action of tumour necrosis factor in mice and rabbits^{8,9}. Other cytokines may also contribute to the pathology of endotoxin shock. But on the basis of the current results we conclude that IL-1 plays an important part in experimental endotoxin shock in animals. It will be worthwhile investigating whether IL-1ra is of practical therapeutic benefit in human septic shock. □

Received 6 July; accepted 23 October 1990.

1. Lode, H. *Arzneimitteltherapie* **1**, 82-89 (1983).
2. Kreger, B. E., Crave, D. E. & McCabe, W. R., *Am. J. Med.* **68**, 343-355 (1980).
3. Guenter, C. A., Florica, V. & Hinshaw, B. J. *J. appl. Physiol.* **26**, 780-786 (1969).
4. Ziegler, E. J., Douglas, H., Sherman, J. E., Davis, C. E. & Braude, A. I. *J. Immun.* **111**, 433-438 (1973).
5. Ziegler, E. J. *et al. New Engl. J. Med.* **307**, 1255-1230 (1982).
6. Ziegler, E., Sprung, C., Straube, R. & Sadoff, J. *Clin. Res.* **38**, 304A (1990).
7. Gorelick, K. J., Schein, R. M. H., MacIntyre, N. R., Emmanuel, G. & Bernard, G. R. *Crit. Care Med.* **18**, S253 (1990).
8. Okusawa, S., Gelfand, J. A., Ikejima, T., Connolly, R. J. & Dinarello, C. A. *J. clin. Invest.* **81**, 1162-1172 (1988).
9. Everaerd, B., Brouckaert, P., Shaw, A. & Fiers, W. *Biochem. biophys. Res. Commun.* **163**, 378-385 (1989).
10. Hannum, C. H. *et al. Nature* **343**, 336-340 (1990).
11. Eisenberg, S. P. *et al. Nature* **343**, 341-346 (1990).
12. Arend, W. P., Welgus, H. G., Thompson, R. C. & Eisenberg, S. P. *J. clin. Invest.* **85**, 1694-1697 (1990).
13. Mancini, G., Carbonara, A. O. & Heremans, J. *Immunochemistry* **2**, 235-254 (1965).
14. Michie, H. R. *et al. New Engl. J. Med.* **318**, 1481-1486 (1988).
15. Tracey, K. J. *et al. Science* **234**, 470-474 (1986).
16. Tracey, K. J. *et al. Nature* **330**, 662-664 (1987).
17. Beutler, B., Milsark, I. & Cerami, A. *Science* **229**, 869-871 (1985).

ACKNOWLEDGEMENTS. This research was supported by the Swedish MRC and the Medical Faculty, University of Lund, Sweden. This study was approved by the Ethics Committee for Animal Experiments at the University of Lund. We thank L. McGee for help in preparing the manuscript.

Phage antibodies: filamentous phage displaying antibody variable domains

John McCafferty*[†], Andrew D. Griffiths*, Greg Winter*[‡] & David J. Chiswell[†]

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[†] Cambridge Antibody Technology Ltd, Daly Research Laboratories, Babraham Hall, Cambridge, CB2 4AT, UK

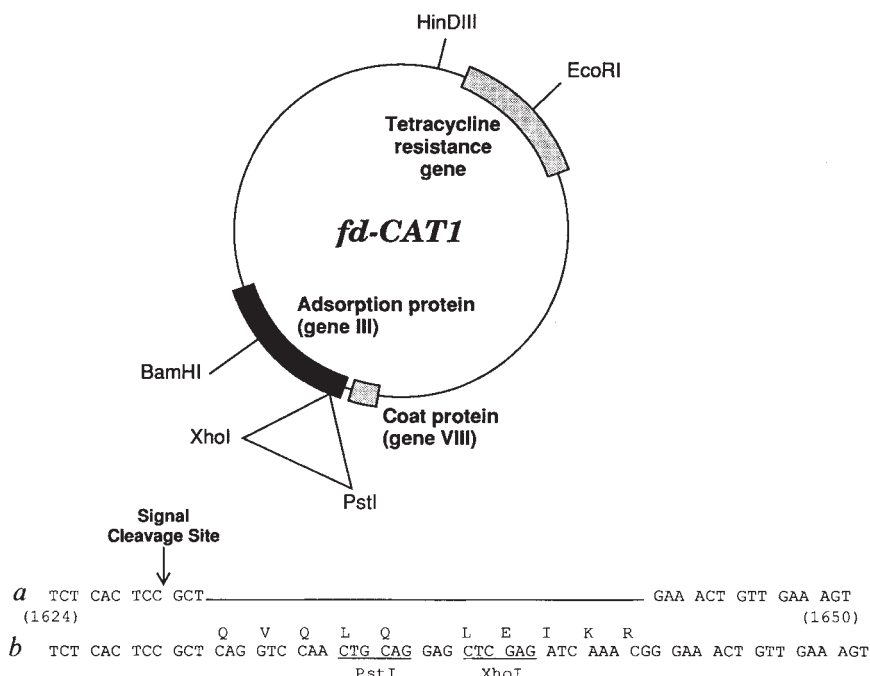
NEW ways of making antibodies have recently been demonstrated using gene technology. Immunoglobulin variable (V) genes are amplified from hybridomas or B cells using the polymerase chain reaction, and cloned into expression vectors. Soluble antibody fragments secreted from bacteria are then screened for binding activities (see ref. 1 for review). Screening of V genes would, however, be revolutionized if they could be expressed on the surface of bacteriophage. Phage carrying V genes that encode binding activities could then be selected directly with antigen. Here we show that complete antibody V domains can be displayed on the surface of fd bacteriophage, that the phage bind specifically to antigen and that rare phage (one in a million) can be isolated after affinity chromatography.

The heavy (VH) and light (VL) chain variable (V) domains of the anti-lysozyme antibody D1.3 (ref. 2) associate tightly as an Fv fragment and bind to antigen with a similar affinity to that of the parent antibody³. To allow expression of both domains on the same polypeptide, they were joined by a flexible linker (Gly₄-Ser)₃ (ref. 4), and the single-chain Fv fragment (scFv) cloned into an fd phage vector (fdCAT1) at the N-terminal region of the gene III protein (Fig. 1). The gene III protein is normally expressed at the tip of fd phage (about four copies per virion), is responsible for attachment of phage to the

[‡] To whom correspondence should be addressed.

FIG. 1 Structure of vector (fdCAT1) for cloning antibody variable domains. Features of the vector used for cloning scFv (D1.3) are shown with *a*, the sequence of fd gene III around the signal peptide cleavage site in fd-tet¹⁸, and *b*, the corresponding region of fdCAT1 with flanking immunoglobulin gene sequences.

METHODS. Vector construction: The vector fd-tet¹⁸ from the American Type Culture Collection was used to transduce *E. coli* strain TG1=K12, (*lac-pro*), *supE*, *thi*, *hsdR5/F' traD36*, *proA* + *B* +, *lacI*^q, *lacZ* M15) to tetracycline resistance, plating onto 2 × YT plates supplemented with 15 µg ml⁻¹ tetracycline¹⁹. First, the *Bst*Ell sites of the vector were removed by digestion, fill-in and religation. To allow the cloning of the scFv (D1.3) gene between the signal sequence and N-terminus of gene III, new restriction sites (*Pst*I and *Xho*I) and flanking immunoglobulin gene sequences were introduced by site-directed mutagenesis²⁰. scFvD1.3 construction. The scFv (D1.3) gene was prepared by digestion of the vector pSW2 (ref. 3) with *Bst*Ell and *Sac*I, and insertion of a synthetic oligonucleotide corresponding to the 3' end of *V_H*, the 5' end of *V_L* and the peptide (Gly₄-Ser)₃ (ref. 4). The scFv was expressed in the bacterial periplasm, shown to bind lysozyme-coated wells by ELISA (data not shown) and then cloned into fdCAT1 at the *Pst*I and *Xho*I sites. Western blots: Concentrated phage (20 µl) (see Fig. 2 legend) was loaded on an SDS-(12.5%) polyacrylamide gel¹¹ and the fractionated proteins transferred electrophoretically to nitrocellulose¹⁰. The filter was blocked in PBS containing 5% skimmed milk powder and 0.1% Tween-20, and developed with rabbit antiserum raised against the D1.3 Fv fragment³ and anti-rabbit antibody conjugated to horseradish peroxidase (Sigma), and detected by enhanced chemiluminescence (Amersham International). The



western blots (not shown) revealed two bands (apparent relative molecular mass (*M_r*) 92,000 and 69,500) suggesting partial proteolysis of the fusion protein. The migration of the upper band was slower than calculated for the fusion protein (*M_r* 69,500), but the migration of gene III protein on SDS-polyacrylamide gels (*M_r* 42,100) is also slower than calculated (apparent *M_r* 55,000–70,000)²¹.

bacterial F pilus⁵, and has been used to display peptide epitopes at the surface of the phage^{6–9}. The antibody-gene III fusion was detected in the recombinant phage (but not in parental fdCAT1 phage) by western blotting¹⁰ of polyacrylamide-SDS gels¹¹, and probing with antisera against the D1.3 Fv fragment³ (data not shown).

Binding of phage to lysozyme was then analysed by enzyme-linked immunosorbent assay (ELISA) (Fig. 2). The phage had the same pattern of reactivity as the D1.3 antibody¹², and bound to hen egg-white lysozyme, but not to turkey egg-white lysozyme, human lysozyme or bovine serum albumin. The specificity of the phage is particularly illustrated by the lack of binding to

the turkey egg-white lysozyme, which differs from hen egg-white lysozyme by only seven amino acids¹². The antigen-binding site is therefore displayed on the surface of the phage and retains antigen binding and specificity. We term phage that display antibody variable domains as phage antibodies.

As phage expressing small peptide epitopes can be purified from mixtures of other phage using antisera or antibodies^{7–8}, we attempted to purify phage antibody (D1.3) using antigen. The phage antibody was mixed with the parental fdCAT1 phage (Fig. 1 legend) and 10¹² phage passed over a column of lysozyme-Sepharose. Colonies derived from the eluates were analysed by probing with an oligonucleotide that detects only

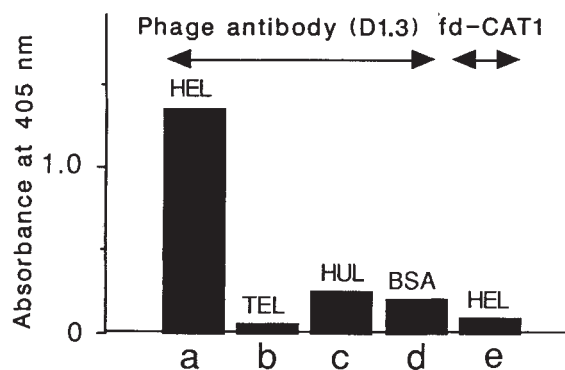


FIG. 2 Binding of phage antibody (D1.3) to lysozymes. Binding of phage as detected by ELISA to hen egg-white lysozyme (*a*, HEL), turkey egg-white lysozyme (*b*, TEL), human lysozyme (*c*, HUL), BSA (*d*). A further control was binding of fdCAT1 to HEL (*e*).

METHODS. Phage growth: Cultures of phage-transduced bacteria were prepared in 10–100 ml 2 × YT medium with 15 µg ml⁻¹ tetracycline and grown with shaking at 37 °C for 16–24 h. Phage supernatant was prepared by centrifugation of the culture (10 min at 10,000 r.p.m. in 8 × 50 ml rotor, Sorval RC-5B centrifuge). At this stage the phage titre was 1–10 × 10¹⁰ ml⁻¹ transducing units. The phage were precipitated by adding 1/5 volume 20% polyethylene glycol, 2.5 M NaCl, leaving for 1 h at 4 °C, and centrifuging (as above). The phage pellets were resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0, to 1/100th of the original volume, and residual bacteria and aggregated phage removed by centrifugation for 2 min in a bench microcentrifuge. ELISA: as described³, except that 2 × 10¹⁰ phage transducing units were added to the antigen-coated plates (1 mg ml⁻¹ antigen) in PBS containing 2% skimmed milk powder. Plates were washed between each step with three rinses of 0.5% Tween-20 in PBS followed by three rinses of PBS. Bound phage was developed by incubating with sheep anti-M13 antisera and detected with horseradish peroxidase conjugated anti-goat serum (Sigma) and ABTS (2,2'-azino-bis(3-ethylbenzthiazoline) sulphonic acid). Absorbance readings were taken at 405 nm after a suitable period.

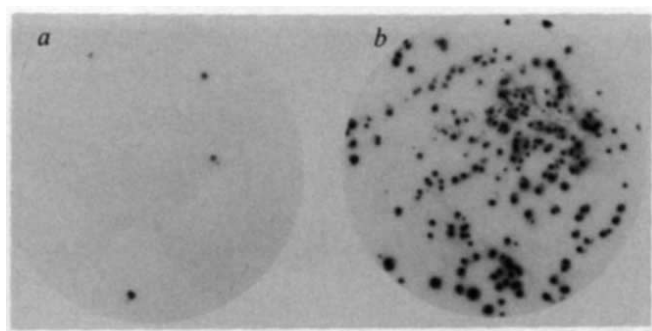


FIG. 3 Oligonucleotide probing of affinity-purified phage. Phage (10^{12}) in the ratio 1 phage antibody (D1.3) in 4×10^4 fdCAT1 phages were affinity-purified and probed with an oligonucleotide specific for phage antibody (D1.3). *a*, Filter after one round of affinity purification (900 colonies total) and, *b*, after two rounds (372 colonies total).

METHODS. Affinity chromatography of phage antibody: About 10^{12} phage particles in 1 ml 2% skimmed milk powder (MPBS) were loaded onto a 1-ml lysozyme-Sepharose affinity column which had been prewashed in MPBS. The column was washed in turn with 10 ml PBS; then 10 ml 50 mM Tris-HCl, 500 mM NaCl, pH 7.5; then 10 ml 50 mM Tris-HCl, 500 mM NaCl, pH 8.5; then 5 ml 50 mM Tris-HCl, 500 mM NaCl, pH 9.5 (adjusted with triethylamine) and then eluted with 5 ml 100 mM triethylamine. The eluate was neutralized with 0.5 M sodium phosphate buffer, pH 6.8, and the phage plated for analysis. For a second round of affinity chromatography, the first column eluate was plated to about 30,000 colonies per Petri dish. After overnight growth, colonies were then scraped into 5 ml $2 \times$ YT medium, and a 20- μ l aliquot diluted into 10 ml fresh medium and grown overnight. The phage was polyethylene-glycol-precipitated as above, resuspended in 1 ml MPBS and loaded onto the column, washed and eluted as above. Oligonucleotide probing. 40 pmol oligonucleotide VH1FOR (ref. 2), specific to phage antibody (D1.3) was phosphorylated and used as probe as in (ref. 19).

the phage antibody (D1.3) (see Table 1 and Fig. 3). At least a thousandfold enrichment of phage antibody (D1.3) was seen with a single column pass. By growing the enriched phage and passing it down the column again, enrichments of up to a millionfold were seen.

Enrichment was also demonstrated using purely immunological criteria. For example, 10^{12} phage (in a ratio 1 phage antibody (D1.3) in 4×10^6 fdCAT1 phage) was subjected to two rounds of affinity selection, and then 26 colonies grown and the phage assayed for lysozyme binding by ELISA. Five colonies yielded phage with binding activities, and these were shown to encode the scFv (D1.3) by polymerase chain reaction (PCR) screening¹³. Thus, very rare phage antibodies can be extracted from large populations by using antigen to select and then screen the phage.

Phage antibodies are likely to find a range of applications in screening *V* genes encoding antigen-binding activities. For example, phage antibodies could be used in cloning and rescue of hybridomas¹⁴, and in the screening of large combinatorial libraries. Thus in the 'random combinatorial' method¹⁵,

TABLE 1 Enrichment of phage antibody (D1.3)

Input ratio*	Output ratio	Enrichment§
Phage antibody: fdCAT1	Oligo† Phage antibody: total phage	ELISA‡ Phage antibody: total phage
Single round		
$1:4 \times 10^3$	43/124	1.3×10^3
$1:4 \times 10^4$	2/82	1.0×10^3
Two rounds		
$1:4 \times 10^4$	197/372	2.1×10^4
$1:4 \times 10^5$	90/356	1.0×10^5
$1:4 \times 10^6$	27/183	5.9×10^5
$1:4 \times 10^7$	13/278	1.8×10^6

* About 10^{12} phage with the stated ratio of phage antibody (D1.3):fdCAT1 were applied to 1 ml lysozyme-Sepharose columns, washed and eluted.

† TG1 cells were infected with the eluted specific binding phage and plated onto $2 \times$ YT plates. After incubation overnight at 30–37 °C the plates were analysed by hybridization to the oligonucleotide VH1FOR (ref. 3) which hybridizes to phage antibody (D1.3) but not to fdCAT1.

‡ Single colonies from overnight plates were grown, phage purified, and tested for lysozyme binding.

§ Enrichment was calculated from the oligonucleotide probing data.

rearranged heavy and light chain *V* genes, isolated by PCR from total splenocytes of a hyperimmunized animal, are combined at random and expressed in *Escherichia coli*. As most of these *V* gene combinations are artificial, those encoding binding activities have a low frequency (1 in 3,000) (ref. 15), and probably lower affinity than original hypermutated VH-VL pairs¹. Rounds of selection using phage antibodies may help in rescuing the higher affinity antibodies. Alternatively, phage antibodies could be used to screen small combinatorial libraries derived from antigen-selected cells¹⁶ to rescue the original VH-VL pairs. Phage antibodies may also be of use in the construction of entirely synthetic antibodies. For example *V* gene repertoires could be made *in vitro* by combining unrearranged *V* genes with *D* and *J* segments¹⁷. Libraries of phage antibodies could then be selected by binding to antigen, hypermutated in the antigen-binding loops *in vitro* and subjected to further rounds of selection and mutagenesis^{1,17}.

The demonstration that a functional antigen-binding site can be expressed on the surface of phage has implications beyond the construction of antibodies. For example, if other protein domains could be expressed at the surface of phage, fd phage vectors could be used to clone and select genes by the binding properties of the expressed protein. Furthermore, endless variants of proteins, including epitope libraries built into the surface of the protein, could be made and readily selected for binding activities. Thus in the future other protein architectures might serve as 'nouvelle' antibodies.

Note added in proof. Insertion of the gene for a scFv directed against the hapten 2-phenyloxazalone into fdCAT1 creates a phage antibody with the predicted specificity. □

Received 29 August; accepted 24 October 1990.

- Winter, G. & Milstein, C. *Nature* (in the press).
- Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Science* **233**, 747–754 (1986).
- Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T. & Winter, G. *Nature* **341**, 544–546 (1989).
- Huston, J. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5879–5883 (1988).
- Kornberg, A. *DNA Replication* (Freeman, San Francisco, 1980).
- Smith, G. *Science* **228**, 1315–1317 (1985).
- Parmley, S. F. & Smith, G. *Gene* **73**, 305–318 (1988).
- Scott, J. K. & Smith, G. *Science* **249**, 386–390 (1990).
- Devlin, J. J., Panganiban, L. C. & Devlin, P. E. *Science* **249**, 404–406 (1990).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Harper, M., Lema, F., Boulot, G. & Poljak, R. J. *Molec. Immun.* **24**, 97–108 (1987).

- Güssow, D. & Clackson, T. *Nucleic Acids Res.* **17**, 4000 (1989).
- Orlandi, R., Güssow, D. H. & Jones, P. T. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 3833–3837 (1989).
- Huse, W. D. *et al. Science* **246**, 1275–1281 (1989).
- Casali, P., Inghirami, G., Nakamura, M., Davies, T. F. & Notkins, A. L. *Science* **234**, 476–479 (1986).
- Milstein, C. *Proc. R. Soc.* **B239**, 1–16 (1990).
- Zacher, A. N., Stock, C. A., Golden, T. W. & Smith, G. P. *Gene* **9**, 127–140 (1980).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1989).
- Taylor, J. W., Ott, J. & Eckstein, F. *Nucleic Acids Res.* **13**, 8764–8785 (1985).
- Goldsmith, M. E. & Konigsberg, W. H. *Biochemistry* **16**, 2686–2694 (1977).

ACKNOWLEDGEMENTS. We thank M. Hobart for the sheep anti-m13 antibody, E. S. Ward for the rabbit anti-Fv D1.3 antisera and C. Milstein for encouragement.

Synthetic analogues of fumagillin that inhibit angiogenesis and suppress tumour growth

Donald Ingber^{*†}, Takeshi Fujita[‡], Shoji Kishimoto[‡], Katsuichi Sudo[‡], Tsuneo Kanamaru[‡], Harold Brem^{*} & Judah Folkman^{*}

^{*}Department of Surgery, The Children's Hospital, and [†]Department of Pathology, Brigham and Women's Hospital, and Departments of Pathology, Surgery, and Anatomy and Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

[‡]Research and Development Division, Takeda Chemical Industries Ltd, Osaka 532, Japan

NEOVASCULARIZATION is critical for the growth of tumours¹⁻³ and is a dominant feature in a variety of angiogenic diseases⁴ such as diabetic retinopathy, haemangiomas, arthritis and psoriasis. Recognition of the potential therapeutic benefit of controlling unabated capillary growth⁵ has led to a search for safe and effective angiogenesis inhibitors. We report here the synthesis of a family of novel inhibitors that are analogues of fumagillin, a naturally secreted antibiotic⁶ of *Aspergillus fumigatus* fresenius. We first isolated this fungus from a contaminated culture of capillary endothelial cells. Purified fumagillin inhibited endothelial cell proliferation *in vitro* and tumour-induced angiogenesis *in vivo*; it also inhibited tumour growth in mice, but prolonged administration was limited because it caused severe weight loss. Synthesis of fumagillin analogues yielded potent angiogenesis inhibitors ('angioinhibins') which suppress the growth of a wide variety of tumours with relatively few side-effects.

During routine culturing of capillary endothelial cells we found a fungal contamination that produced a local gradient of

endothelial cell rounding (Fig. 1). Cells that were only a few cell diameters away were normally spread, suggesting that the fungus was specifically inducing cell rounding rather than causing toxicity. Other fungal contaminants in our experience had produced total cell detachment and cell death. Cell rounding was in itself interesting as endothelial cells must spread in order to proliferate, even when grown in the presence of saturating amounts of soluble mitogens^{7,8}. Also, other angiogenesis inhibitors cause endothelial cell rounding as part of their action *in vivo*⁹.

We isolated this fungus and identified it as *Aspergillus fumigatus* fresenius. Conditioned medium from fungal cultures was found to have potent endothelial cell-rounding activity and to inhibit angiogenesis in the growing chick chorioallantoic membrane. The active fraction was purified from large-scale fungal cultures and identified as fumagillin (Fig. 2), an antibiotic used to treat amoebiasis in humans¹⁰. Purified fumagillin completely inhibited endothelial cell proliferation in the presence of saturating levels of basic fibroblast growth factor (half-maximal inhibition of human umbilical vein endothelial cells at 0.5 ng ml⁻¹). Cell rounding, as in our original contaminated cultures, was induced at concentrations greater than 10 µg ml⁻¹. Angiogenesis was inhibited in the chorioallantoic membrane model at concentrations of fumagillin above 2 µg per chorioallantoic membrane. Fumagillin also suppressed tumour-induced neovascularization in the mouse dorsal air sac (Fig. 3), but its effectiveness as an inhibitor of tumour growth was limited because it produced severe weight loss. This side-effect probably explains the low anti-tumour effect of fumagillin in humans reported previously¹¹.

We therefore set out to design fumagillin analogues that would retain the potent anti-angiogenic activity of fumagillin without its side-effects. Alkaline hydrolysis of fumagillin yields fumagillol, from which over 100 derivatives were synthesized and tested. Among these analogues we identified a subset of compounds that represent a new class of angiostatic antibiotics which we term angioinhibins, one of the most potent of which is *O*-(chloroacetylcarbonyl)fumagillol or AGM-1470 (Fig. 2). This angioinhibin produced half-maximal cytostatic inhibition of endothelial cell proliferation at ~10 pg ml⁻¹, which is 50 times more active than the fumagillin parent. Cytotoxicity (reduction of cell number below the initial plating density) was only observed at much higher concentrations (>1 µg ml⁻¹). Although the sensitivity of a normal non-endothelial cell line (human embryonic lung fibroblasts) was not significantly different from human umbilical vein endothelial cells, doses of AGM-1470 had to be at least ten times higher for comparable inhibition of the

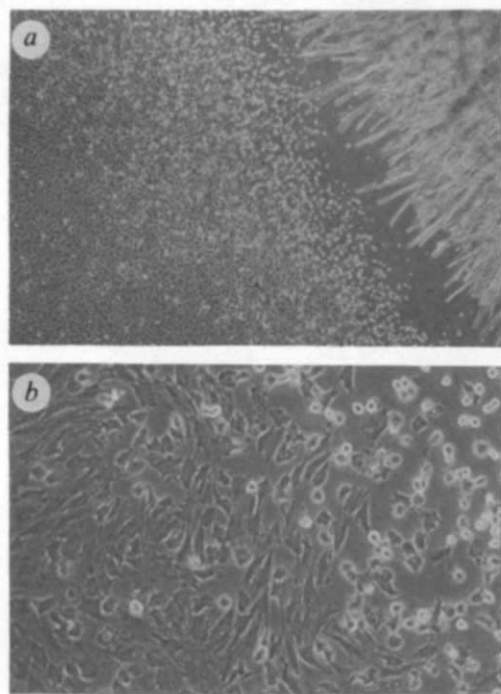


FIG. 1 *a*, A contaminated culture of bovine capillary endothelial cells showing fungal hyphae at the right. *b*, Higher magnification of *a* showing an apparent diffusion gradient resulting in cell detachment and rounding nearest the edge of the fungal colony; cells only a few cell diameters away have normal morphology.

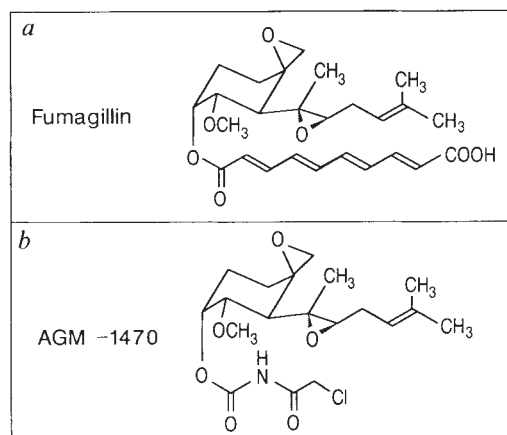


FIG. 2 Structure of *a*, fumagillin, the parent compound isolated from the culture contaminant described in Fig. 1, and *b*, the synthetic analogue AGM-1470, a related angioinhibin used in these antitumour studies.

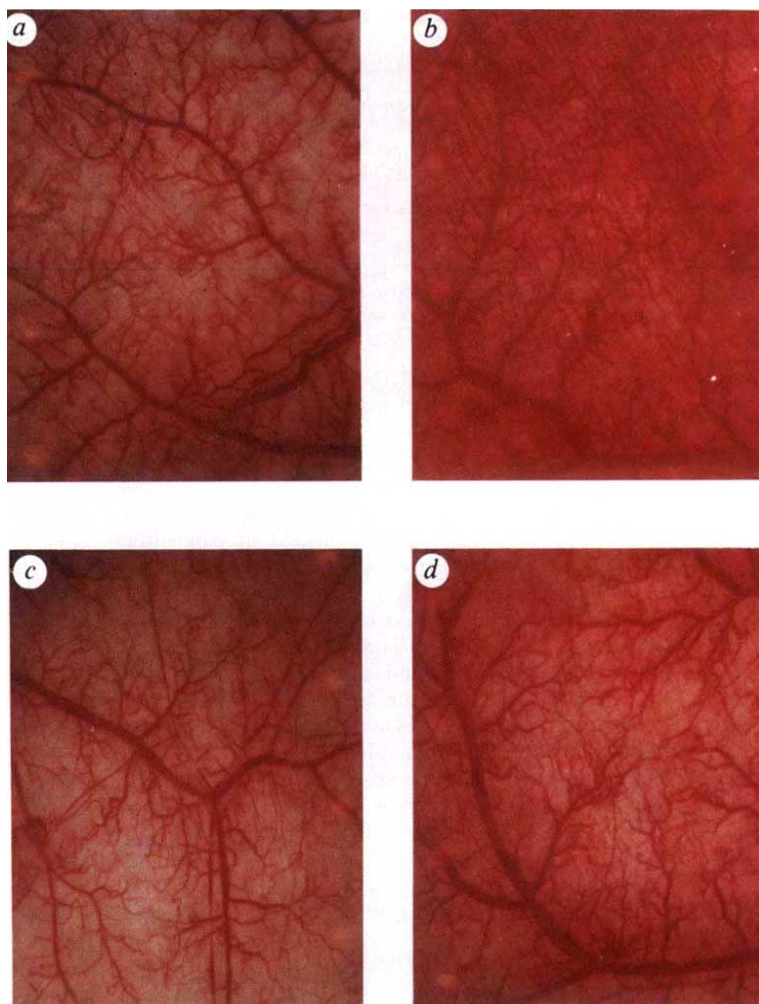


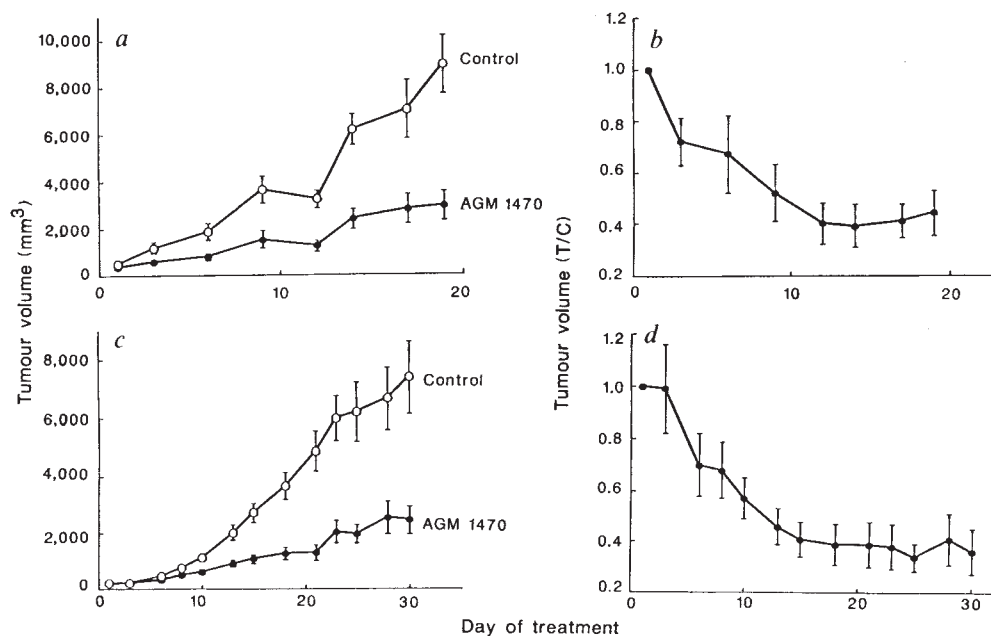
FIG. 3 Inhibition of tumour-induced angiogenesis in the subcutaneous dorsal air sac of the mouse by fumagillin. Chambers containing Millipore filters ($0.45\ \mu\text{m}$) were filled either with saline (*a, c*) or 5×10^6 ascites sarcoma-180 cells in saline (*b, d*) and implanted in dorsal air sacs created surgically in mice. Animals containing the implanted chambers were treated systemically with fumagillin ($100\ \text{mg kg}^{-1}$, subcutaneously) for 3 days (*c, d*). On day 4, angiogenesis within the subcutaneous fascia of fumagillin-treated animals (*c, d*) was compared with that of controls (*a, b*).

human tumour cell lines (squamous cell HSC-1, colon COLO-205, leukaemia HL-60 and squamous cell LS-180).

AGM-1470 also inhibited growth of solid tumours when it was delivered systemically in mice. These tumours included

Lewis lung carcinoma (Fig. 4*a, b*), B16 melanoma (Fig. 4*c, d*), M5076 reticulum cell sarcoma, Meth A sarcoma, colon 26 carcinoma, and Engelbreth-Holm Swarm sarcoma (and others not shown). There was a sustained inhibition of tumour growth

FIG. 4 Inhibition of growth of Lewis lung carcinoma (*a, b*) and B16 melanoma (*c, d*) by systemic administration of AGM-1470. Tumour cells (1×10^6) were inoculated subcutaneously in $0.1\ \text{ml}$ saline in the same position along the dorsal midline in C57B1/6 mice. Treatment was not initiated until tumours were at least 100 to $250\ \text{mm}^3$ in volume. Tumours of similar size were matched for use in control and experimental groups. AGM-1470 was administered subcutaneously at a remote site at $30\ \text{mg per kg}$ in body weight in saline once every other day. Tumours were measured in two dimensions and volume was calculated using a standard formula: $\text{width}^2 \times \text{length} \times 0.52$. T/C = tumour volume of treated/mean tumour volume of control (untreated) for each day. Lewis lung carcinoma, $N=29$ animals per group; B16 melanoma, $N=16$ animals per group. Error bars indicate standard error of the mean.



regardless of whether administration of AGM-1470 was started one day after tumour inoculation (data not shown) or once the tumour had become fully established (when tumour volume exceeded 100 mm³) (Fig. 4). Whereas the application of fumagillin in animals was restricted by their unacceptable weight loss (>15% of starting weight), animals treated with effective doses of AGM-1470 gained weight slowly but consistently. Daily therapy against the tumour was unnecessary in any animal, with less frequent regimens of once every three days, for example, being highly successful (in the case of M5076-bearing animals, survival time increased by up to 260% over untreated controls). Yet cells from the M5076 tumour cell line, which were very sensitive to AGM-1470 *in vivo*, were found to be refractory to AGM 1470 *in vitro*. These tumour cells *in vitro* required 100 times more AGM-1470 than vascular endothelial cells for comparable inhibition of growth. Taken together, these results strongly suggest that AGM-1470 exerts its anti-tumour effects primarily by acting on the tumour vasculature. This hypothesis is further supported by the finding that AGM-1470 did not prolong the survival time of mice injected intraperitoneally with P388 leukaemia. This tumour grows in an ascites (non-solid) form and should therefore be less dependent on neovascularization for its development.

Angiostatics, such as the one reported here, represent a unique class of angiostatic antibiotics. Unlike other angiogenesis inhibitors, angiostatics are not steroid^{12,13}, polysaccharide^{13,14}, retinoid¹⁵ or peptide¹⁵⁻¹⁷ structures. There are two other microbial products with angiostatic activity: one is a sulphated polysaccharide peptidoglycan complex¹⁸ whose complete structure is still unknown; the structure of the other has been determined¹⁹ but it is cytotoxic and its anti-tumour effects have not been reported.

Our new class of angiogenesis inhibitors, then, can suppress tumour growth without the usual toxic side-effects associated with conventional chemotherapy drugs. Specifically, we never observed hair loss, intestinal disturbance or infection although we administered AGM-1470 subcutaneously at 30 mg kg⁻¹ every other day for >100 days (which is about one sixth of the life of a mouse). This absence of side-effects could be critical in treatment of tumours and other angiogenic diseases which may require long-term therapy in much the same way as diabetes and malaria do now. An example of prolonged anti-angiogenic therapy for a non-neoplastic disease has been reported in the treatment of haemangioma with alpha-interferon²⁰. These findings in conjunction with ours provide a glimpse of what a cancer therapy based on anti-angiogenesis might be in the future. □

Received 25 June; accepted 4 October 1990.

1. Folkman, J. *Ann. Int. Med.* **82**, 96-100 (1975).
2. Folkman, J., Watson, K., Ingber, D. & Hanahan, D. *Nature* **339**, 58-61 (1989).
3. Folkman, J. *J. natn. Canc. Inst.* **82**, 4-6 (1990).
4. Folkman, J. & Ingber, D. E. *Ann. Surg.* **206**, 374-383 (1987).
5. Folkman, J. *New Engl. J. Med.* **285**, 1182-1186 (1972).
6. McCowen, M. C., Callender, M. E. & Lawlis Jr, J. F. *Science* **113**, 202-203 (1951).
7. Folkman, J. & Moscona, A. *Nature* **273**, 345-350 (1978).
8. Ingber, D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3579-3583 (1990).
9. Ingber, D., Madri, J. A. & Folkman, J. *Endocrinology* **119**, 1768-1773 (1986).
10. Killough, J. H., Magill, G. B. & Smith, R. C. *Science* **115**, 71-72 (1952).
11. DiPaolo, J. A., Tarbell, D. S. & Moore, G. E. *Antibiotics Annual* 541-546 (1958-1959).
12. Crum, R., Szabo, S. & Folkman, J. *Science* **230**, 1375-1377 (1985).
13. Folkman, J., Langer, R., Lindhart, R., Haudenschild, C. & Taylor, S. *Science* **221**, 719-725 (1983).
14. Folkman, J., Weisz, P. B., Joulie, M. M., Li, W. W. & Ewing, W. R. *Science* **243**, 1490-1493 (1989).
15. Ingber, D. E. & Folkman, J. *Lab. Invest.* **59**, 44-50 (1988).
16. Maione, T. E. *et al. Science* **247**, 77-79 (1990).
17. Moses, M. A., Sudhalter, J. & Langer, R. *Science* **248**, 1408-1410 (1990).
18. Tanaka, N. G. *et al. Canc. Res.* **49**, 6727-6730 (1989).
19. Okawa, T., Hirotsu, K., Shimamura, M., Ashino-Fuse, H. & Iwaguchi, T. *J. Antibiotics* **42**, 1202-1204 (1989).
20. White, C. W., Sondheimer, H. M., Crouch, E. C., Wilson, H. & Fanli, L. L. *New Engl. J. Med.* **320**, 1197-1200 (1989).

ACKNOWLEDGEMENTS. This work was supported by grants from U.S.P.H.S. to J.F. and from Takeda Chemical Industries Ltd to Harvard University. We thank G. Jackson for technical assistance and M. M. Jones for typing the manuscript.

No T-cell tyrosine protein kinase signalling or calcium mobilization after CD4 association with HIV-1 or HIV-1 gp120

Ivan D. Horak*†, Mikulas Popovic‡, Eva M. Horak*, Philip J. Lucas§, Ronald E. Gress§, Carl H. June¶ & Joseph B. Bolen*||

* Laboratory of Tumor Virus Biology, † Medicine Branch, and § Experimental Immunology Branch, National Cancer Institute, Bethesda, Maryland 20892, USA

¶ Naval Medical Research Institute, Bethesda, Maryland 20814, USA

‡ Primate Research Institute, Department of Virology, New Mexico State University, Holloman Air Force Base, New Mexico 88330, USA

THE T lymphocyte surface protein CD4 is an integral membrane glycoprotein noncovalently associated with the tyrosine protein kinase p56^{lck} (reviewed in ref. 1). In normal T cells, surface association of CD4 molecules with other CD4 molecules or other T-cell surface proteins, such as the T-cell antigen receptor, stimulates the activity of the p56^{lck} tyrosine kinase, resulting in the phosphorylation of various cellular proteins at tyrosine residues²⁻⁴. Thus, the signal transduction in T cells generated through the surface engagement of CD4 is similar to that observed for the class of growth factor receptors possessing endogenous tyrosine kinase activity⁵. As CD4 is also the cellular receptor for the human immunodeficiency virus (HIV)^{6,7}, binding of the virus or gp120 (the virus surface protein responsible for specific CD4⁺ T-cell association) could mimic the types of immunological interactions that have previously been found to stimulate p56^{lck} and trigger T-cell activation pathways. We have evaluated this possibility and report here that binding of HIV-1 or the virus glycoprotein gp120 to CD4⁺ human T cells fails to elicit detectable p56^{lck}-dependent tyrosine kinase activation and signalling, alterations in the composition of cellular phosphotyrosine-containing proteins, or changes in intracellular Ca²⁺ concentration.

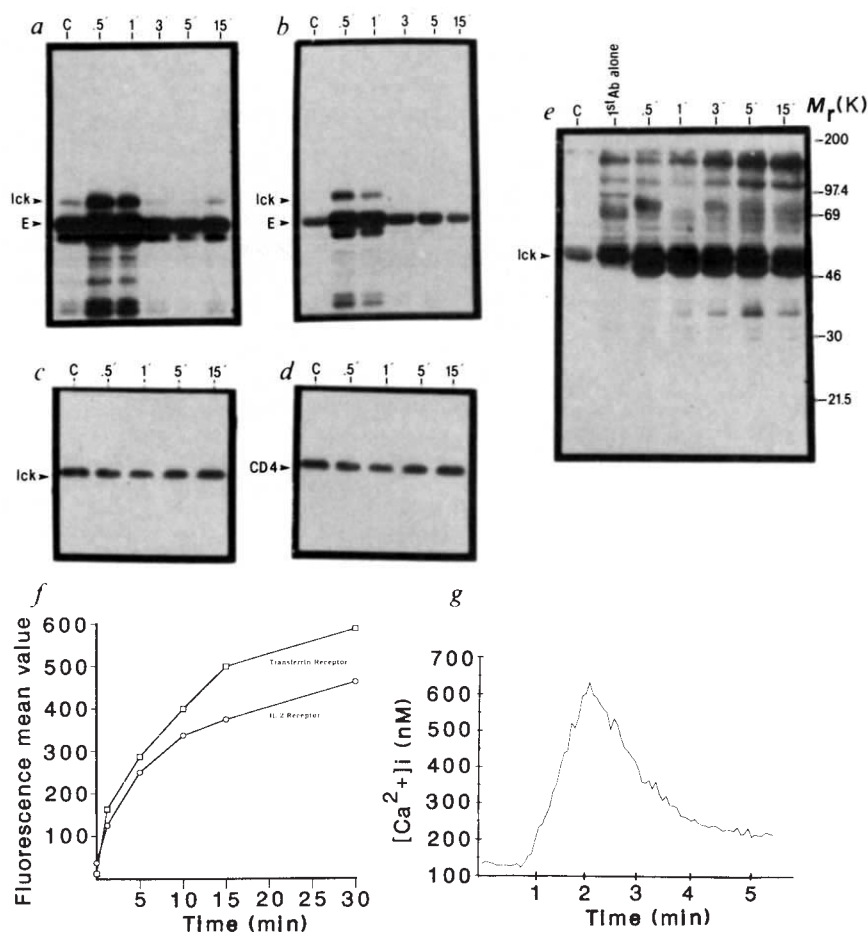
The results presented in Fig. 1 demonstrate that crosslinking surface CD4 molecules with anti-CD4 monoclonal antibodies results in the stimulation of CD4-associated p56^{lck} tyrosine protein kinase activity leading to the phosphorylation of T-cell proteins. Parallel immunoblots detecting p56^{lck} or CD4 show that during the course of these experiments the abundance of p56^{lck}-CD4 complexes was not altered. These results imply that the specific activity of the CD4-associated p56^{lck} was transiently increased around four- to fivefold as a consequence of antibody-mediated CD4 crosslinking. The results of anti-phosphotyrosine immunoblots indicate that cell proteins can be phosphorylated on tyrosine residues in response to the same CD4 crosslinking treatment. In addition, crosslinking of CD4 stimulated Ca²⁺ mobilization after the activation of p56^{lck}, and also led to the rapid appearance of interleukin-2 receptor α subunit (IL-2R α) and transferrin receptor on the surface of the T cells.

To evaluate HIV interactions, purified HIV-1 (strain III-b) was added to the quiescent T cells on ice and binding was monitored by flow cytometry measuring the capacity of the virus to block subsequent binding of fluorescein-conjugated monoclonal antibody Leu3A. HIV-1 could saturate the available surface Leu3A sites within 30 min (Fig. 2a), but these conditions did not result in detectable HIV endocytosis and did not affect the subsequent binding of monoclonal antibody OKT4, which recognizes another CD4 surface epitope distinct from the HIV-binding domain. We found that HIV binding and prolonged

|| To whom correspondence should be addressed at Building 41 Room D-824, National Cancer Institute, Bethesda, Maryland 20892, USA.

FIG. 1 Time course of anti-CD4 mediated cross-linking of CD4 in human T cells. *a*, Immune complex kinase assays following immunoprecipitation with anti-Ick antibodies. *b*, Immune complex kinase assays following immunoprecipitation with anti-CD4 antibodies. The positions of p56^{lck} (Ick) and the exogenous substrate rabbit muscle enolase (E) are shown. *c*, Ick immunoblot of T cell lysates (50 µg protein per lane). *d*, CD4 immunoblot of T-cell lysates (50 µg protein per lane). The position of CD4 is indicated. *e*, Anti-phosphotyrosine immunoblot of T cell lysates (100 µg protein per lane). The position of p56^{lck} and relative molecular mass markers in (K) are indicated. *f*, Fluorescence mean value for the surface detection of IL-2 receptor and transferrin receptor. *g*, The mean [Ca²⁺]_i (nM) using indo-1 loaded T cells.

METHODS. The DP2⁺-allo-specific clone G619-3-53 is a human CD4⁺CD8⁺, IL-2 dependent, helper T cell clone which has been previously described¹⁶. The cells were incubated in the absence of exogenous IL-2 for at least 12 h before use. Cells (3 × 10⁶) were incubated on ice with saturating concentrations (5 µg per 10⁶ cells) of azide-free anti-CD4 monoclonal antibody (Leu3A) (Becton Dickinson) for 30 min, washed and incubated with affinity-purified rabbit anti-mouse IgG (5 µg per 10⁶ cells) (Cappel) at 37 °C for the indicated times before cell lysis (5 × 10⁷ cells per time point). No differences in cellular responses were noted whether Leu3A or monoclonal OKT4 were used in this type of experiment. Immune-complex kinase assays using rabbit muscle enolase as an exogenous substrate; immunoblot analysis using rabbit anti-Ick¹⁷, anti-CD4 (Leu3A), or rabbit anti-phosphotyrosine (Upstate Biotechnology); and flow cytometry measurements of surface antigens (anti-IL-2 receptor and anti-transferrin receptor, Becton Dickinson) have been previously described^{3-4,17}. To measure calcium mobilization, T cells were suspended in HEPES-buffered saline (pH 7.4) containing 1% FCS, loaded with 1.5 µM indo-1 (Molecular Probes) and changes in [Ca²⁺]_i were determined following crosslinking with Leu3A + rabbit anti-mouse IgG



as described above in single cells by flow cytometry as described previously¹⁸.

incubation on ice (up to 1 h) did not elicit detectable changes in p56^{lck} protein kinase activity, the stoichiometry of CD4-p56^{lck} interactions, surface expression of the IL-2R α or transferrin receptor, quantitative or qualitative phosphotyrosine containing cell proteins, nor changes in Ca²⁺ mobilization (data not shown). Figure 2 also demonstrates that raising the temperature to 37 °C after virus absorption also failed to induce detectable changes in any of the parameters previously established as cellular responses following Leu3A-mediated CD4 crosslinking. The same results were obtained with independent HIV-1 preparations and with heterologous peripheral human T cells (data not shown).

We next evaluated the CD4-dependent T-cell signal transduction parameters previously established following binding and crosslinking of purified recombinant HIV-1 gp120 (rgp120). The rgp120 blocked >90% of the binding of Leu3A to T cells within 2 min but did not affect the binding of OKT4 (Fig. 3a). Under conditions allowing for maximal binding rgp120, the CD4-associated rgp120 was crosslinked with pre-titrated goat or rabbit anti-gp120 polyclonal antibodies or with mouse anti-gp120 monoclonal antibody. The results of representative experiments using the mouse anti-gp120 antibodies to crosslink gp120 are shown in Fig. 3. The results demonstrate that p56^{lck} enzyme activity, CD4-p56^{lck} interactions, phosphotyrosine signalling, IL-2R α and transferrin receptor surface expression, and Ca²⁺ mobilization were unchanged after rgp120 crosslinking. The slight upward drift of the mean intracellular Ca²⁺ concentration ([Ca²⁺]_i) plotted in Fig. 3f was not specific for rgp120

crosslinking as the same changes were observed when the cells were incubated at 4 °C and subsequently warmed to 37 °C. Incubation of cells with rgp120 at 37 °C followed by gp120 crosslinking at the same temperature also did not result in detectable changes in [Ca²⁺]_i. Similar results were obtained using goat and rabbit anti-gp120 antibodies as crosslinking reagents, rgp120 from several commercial sources, and gp120 obtained by immunoaffinity purification from HIV-infected H9 cells. Addition of the various anti-gp120 antibodies did not result in dissociation of the bound rgp120 from the cell surface as the Leu3A sites continued to be blocked throughout the course of these experiments. Flow cytometry analysis also indicated that the CD4-rgp120-anti-gp120 immune complexes were retained on the surface of the T cells. Similar results were obtained with peripheral human T cells (data not shown).

We next examined whether gp120 binding and crosslinking could block CD4 signalling mediated by antibody OKT4. Saturating amounts of rgp120 were added to the cells and the bound rgp120 molecules were crosslinked with murine anti-gp120 monoclonal antibody for 15 min. OKT4 antibodies were then added to the same cells and crosslinked by addition of affinity-purified rabbit anti-mouse IgG. The results of this experiment (Fig. 4) show that whereas binding and crosslinking of rgp120 had no apparent effect on the signal parameters previously established, CD4 crosslinking with antibody OKT4 plus rabbit anti-mouse IgG on the same cells resulted in the rapid stimulation of the previously established CD4 responses. These results indicate that binding and crosslinking of rgp120 does

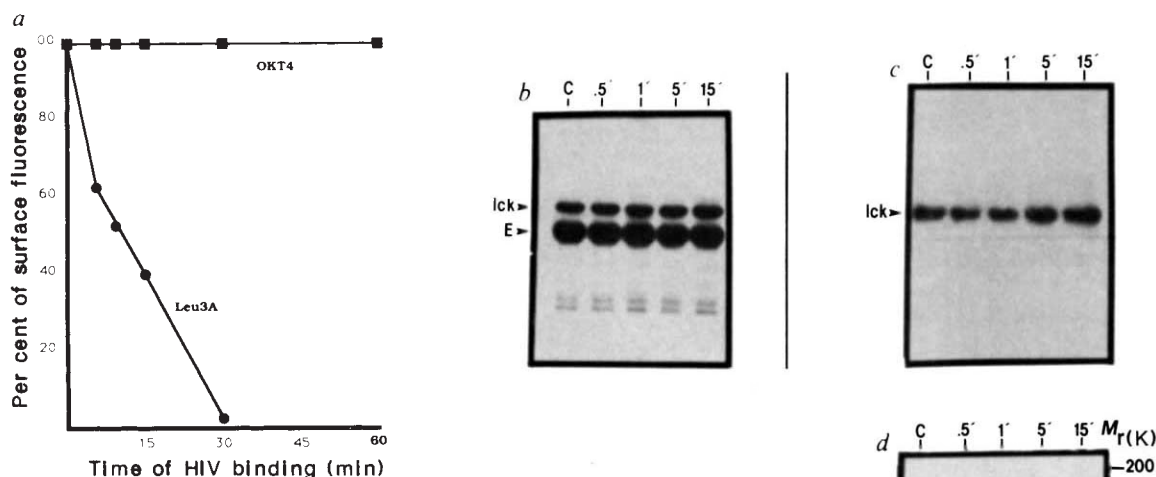
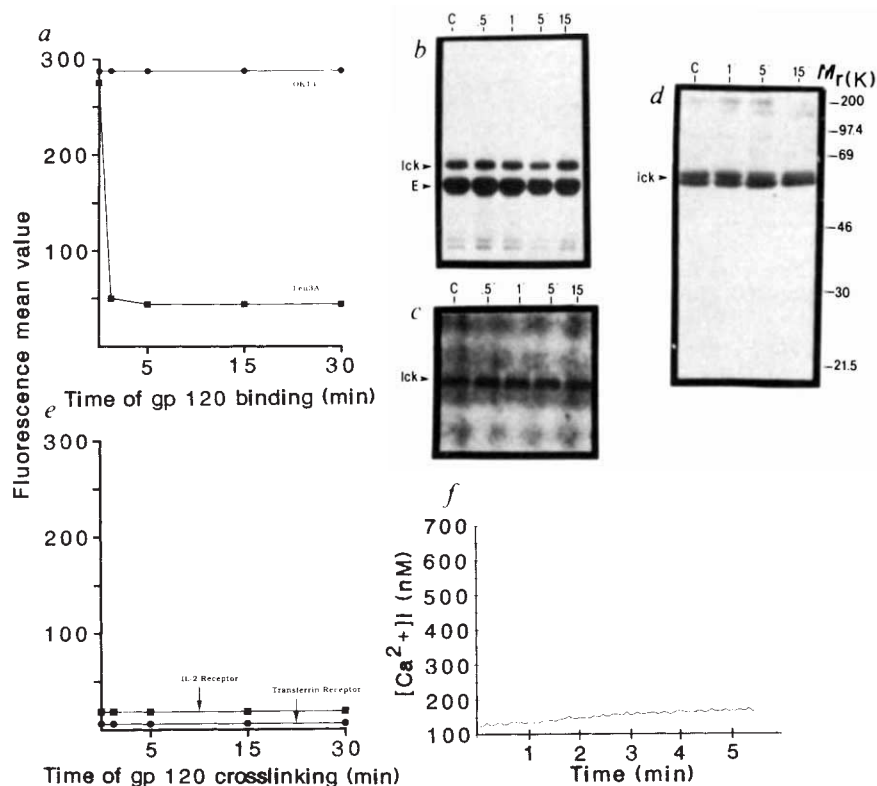


FIG. 2 Effect of HIV-1 binding. *a*, Time course of HIV-1 blocking of Leu3A surface binding sites on CD4. *b*, Immune-complex kinase assays following immunoprecipitation with anti-CD4 antibodies. *c*, Lck immunoblot analysis following immunoprecipitation with anti-CD4 antibodies. *d*, Antiphosphotyrosine immunoblot of T cells lysates.

METHODS. For *a*, clone G916-3-53 cells (3×10^8) were incubated on ice with purified HIV-1 (strain IIIB) for the indicated times. At each point 5×10^6 cells were removed, washed and cold PBS and binding of fluorescein-conjugated monoclonal antibody Leu3A or OKT4 determined by flow cytometry. For the remaining panels, saturating concentrations of HIV-1 were allowed to bind to the cells (3×10^8) for 30 min on ice, cells were washed and warmed to 37 °C for the indicated times. The cells (5×10^7 per time point) were lysed and prepared for immunoprecipitation or immunoblot analysis as previously indicated. To control for potential variations between preparations of purified HIV-1 and to insure that the T cells in each experiment were equally susceptible to HIV-1 infection and capable of virus production, T cells were maintained in parallel cultures in the presence of IL-2 following the binding studies, and HIV-1 reverse transcriptase activity in the culture medium was determined at 8 days (average: 3×10^4 c.p.m. ml⁻¹) and 10 days (average: 1×10^5 c.p.m. ml⁻¹)¹⁹.

FIG. 3 Effects of HIV-1 rgp120 binding and subsequent anti-gp120 mediated crosslinking of CD4. *a*, Time course of HIV-1 rgp120 blocking of Leu3A antibody surface binding sites on CD4. *b*, Immune complex kinase assays following CD4 immunoprecipitation. *c*, Lck immunoblot following CD4 immunoprecipitation. *d*, Anti-phosphotyrosine immunoblot of T cell lysates. The position of p56^{lck} is indicated in *b-d*. *e*, Fluorescence mean value for the surface detection of IL-2 receptor and transferin receptor. *f*, The mean [Ca²⁺]_i (nM) using T cells loaded with indo-1.

METHODS. For *a*, clone G916-3-53 cells (3×10^8) were incubated on ice with purified rgp120 (American biotechnologies) ($10 \mu\text{g}$ per 10^6 cells) for the indicated period of time. At each point 5×10^6 cells were removed, washed with cold PBS and binding of fluorescein-conjugated Leu3A or OKT4 determined by flow cytometry. For the remaining panels, saturating amounts of rgp120 ($10 \mu\text{g}$ per 10^6 cells) were allowed to bind to the cells for 30 min on ice, the cells were washed and the surface rgp120 was crosslinked using anti-HIV-1 gp120 monoclonal antibody (American Biotechnologies) ($5 \mu\text{g}$ antibody per 10^6 cells) at 37 °C. The cells (5×10^7 per time point) were then collected at the indicated times and the various assays conducted as described in Fig. 1.



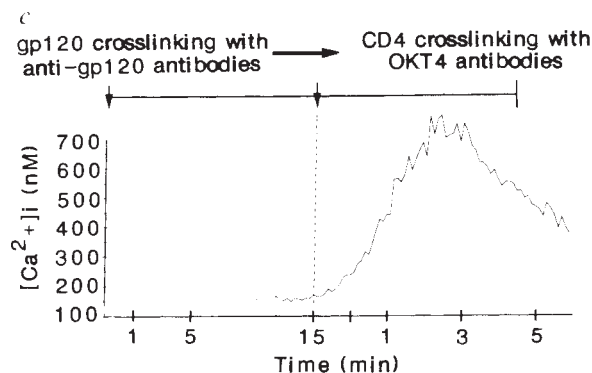
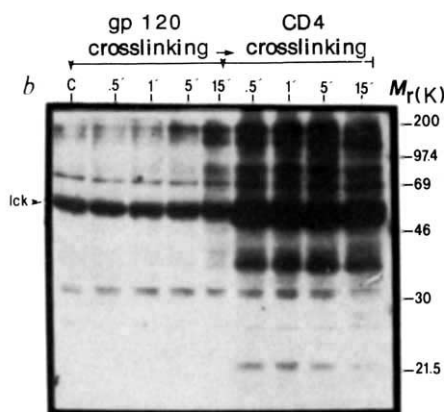
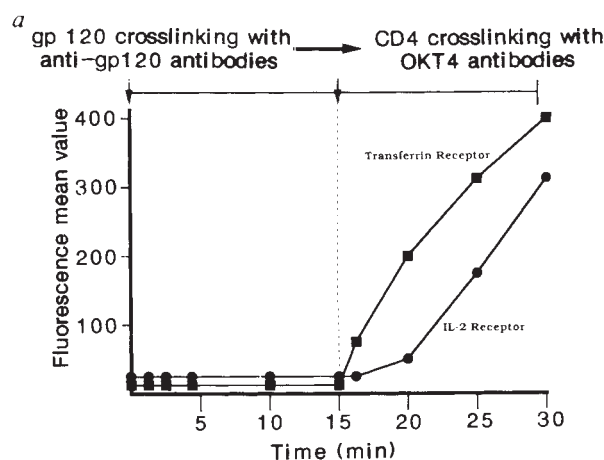


FIG. 4 Effect of HIV-1 gp120 binding on crosslinking of CD4 by anti-CD4 Mab. *a*, Time course of IL-2 receptor and transferrin receptor surface expression accompanying HIV-1 gp120 immunological crosslinking and following subsequent CD4 immunological crosslinking with OKT4. *b*, Parallel antiphosphotyrosine immunoblot of the T-cell lysates. *c*, Parallel determination of the mean $[Ca^{2+}]_i$.

METHODS. Clone G916-3-53 cells were incubated with rgp120 and the rgp120 immunologically crosslinked at 37 °C for the indicated times. The cells were then incubated on ice with OKT4 (1.5 μ g per 10^6 cells) for 30 min followed by immunological crosslinking with rabbit anti-mouse IgG (5 μ g per 10^6 cells) as described in Figure 1.

not inhibit CD4-mediated signal transduction events stimulated following CD4-CD4 interactions.

Our results show that absorption of HIV-1 or binding and crosslinking of HIV-1 gp120 does not stimulate the human CD4⁺ T lymphocyte signal transduction pathways that are triggered upon surface CD4 engagement mediated by anti-CD4 antibodies. These results support the finding that non-proliferating T cells can be infected with HIV⁸⁻¹¹ and indicate that early HIV-1 interactions with quiescent T cells may not generate the types of stimulatory signals that have been established as requirements for virus production. Although in agreement with the data of Mittler and Hoffman¹², our results differ from other reports demonstrating that CD4-dependent gp120 binding to T cells^{13,14} results in increased Ca^{2+} mobilization and stimulation of other intracellular activation signals. Increased Ca^{2+} mobilization has also been reported following CD4-independent gp120 binding to cultured rodent retinal ganglion cell neurons¹⁵. An explanation for the discrepancies is not clear at present. □

Received 31 May; accepted 21 September 1990.

1. Bolen, J. B. & Veillette, A. *Trends biochem. Sci.* **14**, 404-407 (1989).
2. Veillette, A., Bookman, M. A., Horak, E. M., Samelson, L. E. & Bolen, J. B. *Nature* **338**, 257-259 (1989).
3. Veillette, A., Zuniga-Pflucker, J., Bolen, J. B. & Kruisbeek, A. M. *J. exp. Med.* **170**, 1671-1680 (1989).
4. Veillette, A., Bolen, J. B. & Bookman, M. A. *Molec. cell. Biol.* **9**, 4441-4446 (1989).
5. Ullrich, A. & Schlessinger, J. *Cell* **61**, 2-3-212 (1990).
6. Sattentau, Q. J. & Weiss, R. *Cell* **52**, 631-633 (1988).
7. Robey, E. & Axel, R. *Cell* **60**, 697-700 (1990).
8. Folks, T. et al. *J. Immun.* **136**, 4049-4053 (1986).

9. Zagury, D. et al. *Science* **231**, 850-853 (1986).
10. Zack, J. A., Cann, A. J., Lugo, J. P. & Chen, I. S. Y. *Science* **240**, 1026-1029 (1988).
11. Zack, J. A. et al. *Cell* **61**, 213-222 (1990).
12. Mittler, R. S. & Hoffmann, M. K. *Science* **245**, 1380-1382 (1989).
13. Kornfeld, H., Cruikshank, W. W., Pyle, S. W., Berman, J. S. & Center, D. M. *Nature* **335**, 445-448 (1988).
14. Neudorf, S. M., Jones, M. M., McCarthy, B. M., Harmony, J. A. & Choi, E. M. *Cell. Immun.* **125**, 301-314 (1990).
15. Dreyer, E. B., Kaiser, P. K., Offermann, J. T. & Lipton, S. A. *Science* **248**, 364-367.
16. Golding, H. et al. *J. clin. Invest.* **83**, 1430-1435 (1989).
17. Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. *Cell* **55**, 301-308 (1988).
18. Rabinovitch, P. S., June, C. H., Grossmann, A. & Ledbetter, J. A. *J. Immun.* **137**, 952-961 (1986).
19. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).

ERRATUM

Targets of homeotic gene control in *Drosophila*

Alex P. Gould, Jenny J. Brookman, David I. Strutt & Robert A. H. White

Nature **348**, 308-312 (1990)

In the above article, the first line of the immunopurification strategy shown in Fig. 1 should have read: Nuclei from 3-15 h embryos.

Today's SCID-hu mouse

H. Kaneshima, C. Baum, B. Chen, R. Namikawa, H. Outzen, L. Rabin, A. Tsukamoto and J.M. McCune

Use of the SCID-hu mouse — an immunodeficient mouse transplanted with relevant organs of the human immune system — as a preclinical testing system for human therapeutics is discussed.

THE haematolymphoid system of the laboratory mouse has been extensively explored. Nevertheless, it is difficult to extrapolate findings made in the mouse to events as they occur in man. First, one cannot be certain that concordant physiological and pathophysiological attributes will be shared and, second, validity cannot be established without embarking upon time-consuming and costly human clinical trials.

A relevant animal model for human haematopoiesis may help circumvent some of these obstacles. The development of the SCID-hu mouse, a heterochimaeric mouse developed by surgical transplantation of interactive human haematolymphoid organs into the immunodeficient C.B-17 *scid/scid* mouse has been described¹. Engraftment and partial reconstitution of the mouse with human lymphoid organs and the cells within them was documented. It was determined that human B cells, T cells and antigen-presenting cells appeared to function normally, at least within the context of engrafted human lymph nodes. Some of the improvements that have since been made in the 'construction' and use of SCID-hu mice are discussed in this article (see Fig. 1), but more detailed information can be found in several reviews²⁻⁴.

HIV-1 infectivity assay

The SCID-hu mouse has been found to be permissive for replication and spread of human immunodeficiency virus type 1 (HIV-1) after direct injection of virus into the engrafted human lymph node or thymus⁵. Signs of infection are evident in a time- and dose-dependent manner and viral replication is suppressed if the antiviral compound 3'-azido-3'-thymidine (AZT) is administered before inoculation of virus⁶. The host range of HIV-1 in the SCID-hu mouse is similar to that found in humans: CD4⁺ T cells and myelomonocy-

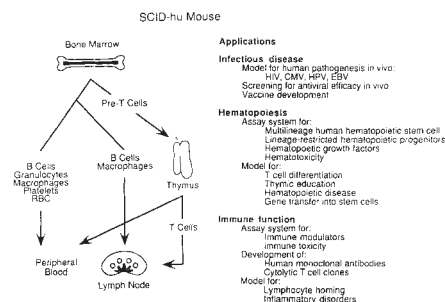


FIG. 1 Potential applications of the SCID-hu mouse.

tic cells are infected, murine cells in the endogenous mouse organs are not^{7,8}. The relevance of this system to the analysis of HIV infection in man is best underscored by the observation that only primary clinical isolates of HIV-1 can replicate in the human organs of the SCID-hu mouse⁹. In contrast, isolates that have been adapted to tissue culture growth are not infectious in this *in vivo* setting.

Recently, it was shown that intravenous infection of SCID-hu mice can result in viraemia^{8,10}. This observation has facilitated the development of a reproducible and practical *in vivo* infectivity assay for HIV. Within two weeks of intravenous injection of a standard stock of virus, the human lymphoid organs in all test animals are infected and, of these, 95 per cent are viraemic. This assay can be used for the evaluation of compounds *in vivo*. Questions about bioavailability and efficacy can now be asked that could otherwise be answered only in human clinical trials. For example, when AZT and 2', 3'-dideoxyinosine were tested, a 3-4-week assay in the SCID-hu mouse generated dose-response data equivalent to that now being used in man^{8,11}. In an experiment that would be difficult to carry out in man, post-exposure administration of AZT to the HIV-infected SCID-hu mouse has been shown to suppress viral

replication in a time-dependent manner¹².

Analysis of haematopoiesis

Early constructions of SCID-hu mice provided a microenvironment (the human fetal thymus) that was supportive of human T cell and myeloid differentiation¹. Newly developed techniques have since provided substantial qualitative and quantitative improvements. Simple operative procedures have been devised for the co-implantation of human fetal liver and thymus. In many cases, a conjoint organ forms with areas of extramedullary haematopoiesis that closely resemble normal human bone marrow¹³. In over 50 per cent of such 'Thy/Liv' mice, multilineage human haematopoiesis is observed *in vivo* for as long as 5-11 months. The microenvironmental space of the human thymus graft has also been used as a read-out for purified populations of haematopoietic cells¹⁴. After microinjection, only a limited cell population demonstrates multilineage differential potential. In this way, the SCID-hu mouse is proving to be instrumental in the isolation and definition of the human haematopoietic stem cell.

Analysis of immune functions

In recent studies of the reconstitution of human immune function within the SCID-hu mouse, it was found that human T cells in cohorts of Thy/Liv mice circulate with a normal CD4/CD8 ratio and do not express the activation antigens, CD25 and CD69; these findings are consistent with the apparent absence of graft-versus-host disease¹⁵. The human T cells of the SCID-hu mouse do, however, respond to mitogens, antibodies against CD3, and alloantigens. Consequently, it is now possible to analyse the effects of extraneous agents, such as HIV, on the function of such cells *in vivo*.

Techniques have also been developed for the construction of SCID-hu mice that support defined, primary human immune responses. Animals have been immunized with test antigens, such as keyhole limpet haemocyanin conjugated with trinitrophenol (KLH-TNP), and found to generate specific human antibody responses against TNP (and KLH). The human B-lineage cells involved in these responses can be isolated from the SCID-hu mouse and immortalized to produce human monoclonal antibodies. This is similar to

1. McCune, J.M. *et al.* *Science* **241**, 1632-1639 (1988).
2. McCune, J.M., Kaneshima, H., Lieberman, M., Weissman, I.L. & Namikawa, R. in *Curr. Topics Microbiol. Immunol.* Vol. 152 (eds Bosma, M.J., Phillips, R.A. & Schuler, W.) 183-193 (Springer, Berlin, 1989).
3. McCune, J.M. in *Semin. Virol.* Vol. 1 (ed. Chen, I.S.Y.) 229-235 (W.B. Saunders, Philadelphia, 1990).
4. McCune, J.M. in *Human Retroviruses* Vol. 119 (eds Groopman, J.E., Chen, I.S.Y., Essex, M. & Weiss, R.A.) 347-359 (Wiley, New York, 1990).
5. Namikawa, R., Kaneshima, H., Lieberman, M., Weissman, I.L. & McCune, J.M. *Science* **242**, 1684-1686 (1988).
6. McCune, J.M., Namikawa, R., Shih, C.-C., Rabin, L. & Kaneshima, H. *Science* **247**, 564-566 (1990).
7. Namikawa, R., Fedor, J., Kaneshima, H. & McCune, J.M. in *6th int. Conf. AIDS* Vol. 1, 194 (1990).

8. Kaneshima, H., Shih, C.-C., Namikawa, R., Rabin, L. & McCune, J.M. (submitted).
9. Rabin, L., Kaneshima, H., Shih, C.-C. & McCune, J.M. in *6th int. Conf. AIDS* Vol. 3, 113 (1990).
10. Kaneshima, H., Shih, C.-C., Namikawa, R., Rabin, L. & McCune, J.M. in *6th int. Conf. AIDS* Vol. 3, 99 (1990).
11. McCune, J.M., Shih, C.-C., Namikawa, R., Rabin, L. & Kaneshima, H. in *6th int. Conf. AIDS* Vol. 2, 107 (1990).
12. Shih, C.-C. *et al.* *J. infect. Dis.* (in the press).
13. Namikawa, R., Weilbaecher, K.N., Kaneshima, H., Yee, E.J. & McCune, J.M. *J. exp. Med.* **172**, 1055-1063 (1990).
14. Peault, B., Weissman, I.L., Baum, C., McCune, J.M. & Tsukamoto, A. (submitted).
15. Krowka, J., Sarin, S., Namikawa, R., McCune, J.M. & Kaneshima, H. in *6th int. Conf. AIDS* Vol. 1, 194 (1990).

the techniques used to generate murine monoclonal antibodies except, that after immunization, the human organs are harvested instead of the mouse spleen. Future refinements in this technology may make it possible to generate human monoclonal antibodies against a variety of medically important antigens.

The SCID-hu mouse has come a long way in developmental terms over the past two years. Today, 'minimal essential mice' are specifically designed for the analysis of medically important problems that are difficult, if not impossible, to study in man. In this way, the SCID-hu mouse may serve both as a tool for the selection of important modalities (for example, antiviral compounds or human haematopoietic progenitor cells), and as a vehicle for their development (for example, human monoclonal antibodies). □

Hidetoshi Kaneshima, Charles Baum, Benjamin Chen, Reiko Namikawa, Henry Outzen, Linda Rabin, Ann Tsukamoto and Joseph M. McCune are at SyStemix, Inc., 3400 West Bayshore Road, Palo Alto, California 94303, USA. The authors would like to acknowledge John Krowka, Jan Laman, Susan Mayo, Bruno Peault, Chu-Chih Shih and Phil Streeter at SyStemix, and Irving L. Weissman at Stanford University for their useful suggestions. We thank the Transplantation and Animal production groups for their help in making these studies possible. For more information, fill in reader service number 100.

ADVERTISEMENT

Reader Service No.25

Teflon[®] Coated and Bare Wire from MEDWIRE



Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E.I. DuPont & Co.

Reader Service No.26

Custom DNA

Purified and Shipped in 48 hours, Worldwide.

100 ug. \$10.00 a base for the first 15 bases, \$7.50 for each additional base.

Delivered.

Research Genetics

1-800-533-4363

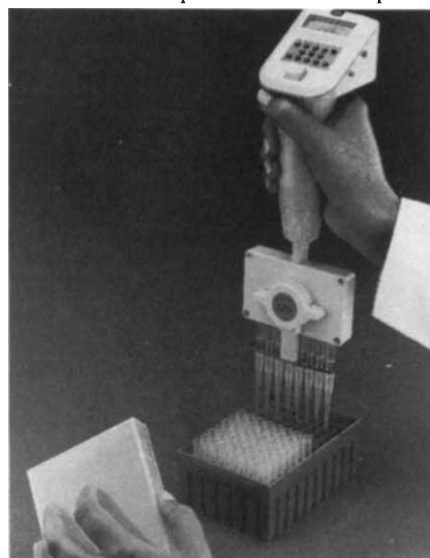
In the United Kingdom 0-800-89-1393
In Canada 1-800-533-4363

FAX 205-536-9016

Cell science

Next week's conference sponsored by the American Society for Biological Chemistry and Molecular Biology/American Society for Cell Biology will be held in San Diego, California. Year-end products include a cDNA synthesis kit and an image processing system that offers a more affordable approach to confocal microscopy.

EDP-Plus M8 is an eight-channel, 250 µl motorized electronic pipette from Rainin designed for use with 96-well plates (Reader Service No. 101). The pipette has ten pipetting speeds and five liquid measurement modes: pipette, multiple dispense, dilute, titrate and measure. In addition, EDP-Plus M8 can be used to mix liquids after dispensing and to pick up and deliver up to 12 different volumes in sequence, says Rainin. To ensure repeatable liquid measurement that is independent of user technique, the EDP-Plus M8 uses computer-controlled piston



EDP-Plus M8 — Rainin's motorized and portable eight-channel pipette.

movements. The pipette also uses a rechargeable battery for complete portability. EDP-Plus M8 sells for about \$800 (US) and can be seen in action in booth 710.

Oncogene Science is offering two quantitative immunoassays for human epidermal growth factor (EGF) and granulocyte colony stimulating factor (G-CSF), which are based on a sandwich ELISA and provide results within six and eight hours, respectively (Reader Service No. 102). The G-CSF assay is a two-site immunometric assay that uses specific anti-G-CSF monoclonal antibodies, pre-coated onto a 96-well microtitre plate, to capture G-CSF antigen in the test sample. Bound G-CSF is then incubated with a biotinylated polyclonal G-CSF antibody.



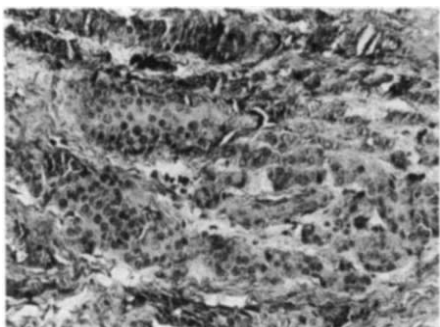
Immunoassays for EGF and G-CSF.

This 'reporter' antibody recognises the G-CSF retained by the 'capture' antibody and is detected, in turn, by streptavidin-horseradish peroxidase and substrate. The absorbance is measured and the amount of G-CSF determined from a standard curve. The principle of the EGF assay is the same. The immunoassays can detect G-CSF and EGF — both native and recombinant human forms — in serum, plasma and tissue culture media with a sensitivity of 50 pg ml⁻¹ and 10 pg ml⁻¹, respectively. The \$475 (US) kits include pre-coated 96-well plates, standards, biotinylated affinity-purified antibodies, buffers and instructions. Oncogene Science's products will be on display in booth 1328.

Monoclonal assortment

Through an agreement with Serotec, Bioproducts for Science is distributing a new mouse monoclonal antibody to epidermal growth factor (EGF) (Reader Service No. 103). Potential applications for the antibody include RIAs, ELISAs, immunoprecipitation, immunoblotting and immunopurification. The EGF antibody reacts 100 per cent with human EGF — both recombinant and native forms — and about five per cent with natural mouse EGF, says BPS. It neutralizes EGF activity *in vitro* and inhibits binding of ¹²⁵I-EGF to the EGF receptor. The antibody is supplied as a 2-ml tissue culture supernatant for \$276.82 (US). BPS will be in booth 1538.

Anti-proliferation cell nuclear antigen (PCNA) — also known as cyclin — is the latest Path-Mark/Coulter Clone monoclonal antibody (mAb) to roll off the Coul-



Anti-PCNA monoclonal (cyclin), intended for immunohistochemical research applications. ter production line (*Reader Service No. 104*). The new **anti-PCNA monoclonal** can be used with paraffin and frozen tissue sections for immunohistochemical research applications. According to Coulter, recent studies have shown a direct correlation between PCNA content and the rate of DNA synthesis and cellular proliferation. It has also been reported that the expression of PCNA may be used as early indicator of certain malignancies. Other products in the Path-Mark/Coulter Clone mAb line include anti-leukocyte common antigen, anti-alpha fetoprotein, anti-carcinoembryonic antigen and anti-prostate specific antigen. All are supplied in a purified and pre-diluted form in colour-coded dropper bottles. To complement the mAbs, Coulter offers two immunoperoxidase kits: the universal Streptavidin Immunoperoxidase Kit and the universal ABC Immunoperoxidase kit. See Coulter's Scientific Instruments division in booth 1015.

Caltag Laboratories, exhibiting in booth 731, has developed a **murine monoclonal antibody to P-Glycoprotein** called clone JSB-1 (*Reader Service No. 105*). JSB-1 was obtained by immunizing a BALB-c mouse with multi-drug resistant (MDR) Chinese hamster ovary cell line CHRC5. The antibody reacts with a conserved cytoplasmic epitope of the plasma-membrane associated 170–180-K glycoprotein, the expression of which is correlated with the degree of multi-drug-derived MDR cell lines and human MDR cell lines, says Caltag. This includes cell lines derived from lung, ovaries and B-cell lymphomas. The antibody can be used with air-dried or acetone-fixed cells, as well as acetone-fixed cryostat sections, and is available as prediluted ascites fluid in 1-ml (100–200 tests) and 5-ml (500–1,000 tests) quantities for \$215 and \$900 (US), respectively. For more on Caltag's monoclonals — see booth 731.

DNA/RNA details

All the reagents you need to **generate first-strand cDNA from total or poly(A)⁺ RNA** are contained within the SuperScript Preamplification System from Life Technologies (*Reader Service No. 106*). When used in conjunction with

polymerase chain reaction technology (PCR), the SuperScript System can, says Gibco/BRL be used to detect the presence of rare messages, to quantitate the amount of specific mRNA from small numbers of cells, or to clone specific cDNAs, without having to construct an entire cDNA library. The system uses RNase H⁻ reverse transcriptase, which has been engineered to eliminate RNase H activity. Following first strand synthesis, a controlled RNase H digestion removes the RNA template; this improves PCR amplification, says the company. In addition, as the reverse transcriptase is not inhibited by ribosomal or transfer RNA, either total or oligo(dT) purified RNA can be used. The SuperScript System includes sufficient reagents to perform 50 first-strand cDNA synthesis reactions. Gibco/BRL also supplies a control RNA and two control primers for monitoring performance and a manual containing a simplified RNA isolation procedure. See booth 1614.

Promega offers three new strains of **competant cells** for standard subcloning purposes or for the generation of cDNA libraries (*Reader Service No. 107*). The high- and low-efficiency strains are HB101, JM109 and MB408. All cells are supplied in 0.2-ml aliquots. In addition, Promega supplies 3 ng of pGEM-3Z plasmid DNA as a test control vector. Stop by the company's booth, number 637.

Labware line-up

Protect your temperature-sensitive enzymes from temperature fluctuations caused by open freezer doors, power failures and freeze/thaw cycles with the StrataCooler Lite **bench-top cooler unit** from Stratagene (*Reader Service No. 108*). Measuring 9 cm in diameter by 9 cm



in height, the lightweight StrataCooler unit has been shown to maintain a temperature of -10°C for up to one hour when outside the freezer, says Stratagene. The \$75 (US) StrataCooler unit can be used to store seven tubes in 1.15-cm diameter holes that are 4 cm deep.

Eppendorf has introduced new **software** for the company's thermal cycler — the MicroCycler E (*Reader Service No. 109*). Version 1.01 provides a real-time clock that allows the instrument to predict completion times with increased accu-



Eppendorf's thermal cycler features a software-upgrade — version 1.01.

acy, says Eppendorf. In this way, users can schedule experiments for unattended operation. The MicroCycler E, which is available in 0.5- and 1.5-ml models, features menu-based programming, which offers complete control of temperature profiles. The sample temperature sensor and linear block design ensure well-to-well uniformity and temperature accuracy, says Eppendorf. Information about run status, program steps and cycles, and multiple program lines can be viewed on the LED. Existing owners of a MicroCycler E will receive the software upgrade free of charge.

Chemiluminescent-based kits

ChemiProbe, the complete **nonradioactive labelling system** for DNA or RNA probes and the detection of these probes following hybridization, is now available from Bios (*Reader Service No. 110*). The kit is based on the incorporation of a sulphone label into the cytosine residues of the DNA or RNA. The modified DNA site then serves as an antibody receptor. In the first step of a two-step immunoenzymatic reaction that follows, a monoclonal antibody is bound to the sulphone label, which then serves as the binding site for an enzyme-antibody conjugate. The enzyme causes a chromogenic substrate to form a semi-permanent, coloured precipitate. Stated applications include *in situ* hybridization, Southern blotting, dot blotting, colony hybridization and plaque hybridization. With ChemiProbe, results can be achieved in less than two hours, with a sensitivity of 0.2 pg in Southern blot procedures, says Bios. Using this set-up, large amounts of probe DNA can be labelled at once and stored for up to one year. Bios' booth is 1305.

For the **non-isotopic detection of proteins and nucleic acids** using chemiluminescence, United States Biochemical recommends the Protein Images western blotting detection kit (*Reader Service No. 111*). The detection chemistry involved provides a safe and sensitive alternative to radioactivity and is based on Lumi-Phos 480, a chemiluminescent substrate for alkaline phosphatase. The substrate is stable indefinitely when stored at 4°C , says USB. The PVDF membranes used in

ADVERTISEMENTS

LIPOSOMES
REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM
POB 126, D-8000 Munich 65, FRG

Reader Service No.18

Synthetic Oligonucleotides
Custom made Probes and Primers,
Synthesised, purified and shipped in 3 days

£4.00 per base
(All inclusive)
0.2 µM scale (approx. 500 µg)
Modified bases, aminoalkylation,
data base analysis and full design facilities
Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX
Tel. XX-44-(0)71 - 326 3126
FAX XX-44-(0)71 - 733 3877

Reader Service No.56

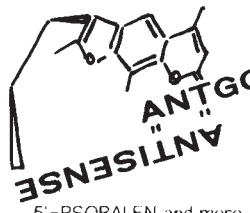
TIB MOLBIOL BERLIN
CUSTOM DNA SYNTHESIS

5,- DM per base

DM 95,- per oligo
FAX (0049) 30 6261960

Reader Service No.79

CUSTOM DNA SYNTHESIS



ANTGCCAT
ANTISENSE

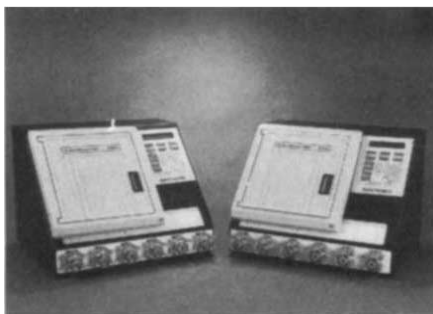
5'-PSORALEN and more.....
TIB MOLBIOL BERLIN
FAX (0049) 30 6261960

Reader Service No.19

the kit are designed to provide low background with high sensitivity. Blots can be re-probed and re-exposed. The \$275 (US) Protein Images kit includes membranes, blotting paper, heat-sealable plastic bags for incubating and washing the blots, and a choice of three commonly used alkaline phosphatase-conjugated antibodies. USB's range of non-isotopic kits can be seen in booth 1331.

Cell culture ware

Two new **automated mammalian cell culture systems** — the Acusyst-Maximizer 500 and 1000 — are available from Endotronics and (Reader Service No. 112). The two systems differ in their level of productivity: with the Acusyst-Maximizer 500 (\$32,500 (US)) and 1000 (\$42,500 (US)), typical hybridoma cells will produce 500 and 1,000 mg per day, respectively, says Endotronics. The systems are



Maximize your yields with the Acusyst-Maximizer mammalian cell culture system.

designed to support a wide variety of cell lines and use pre-sterilized assembled cultureware containing a single or double hollow-fibre bioreactor with 10,000 Da cutoff. In this way, users can minimize the quantity of supplements required and maximize the concentration and purity of high molecular weight harvested products, says the company. To meet specific needs, Endotronics will supply customized cultureware for applications requiring high molecular weight or microporous membranes. The system monitors dissolved oxygen, and controls pH, temperature, circulation rates, and six input/output pumps. The Acusyst-Maximizer will be on display in booth 506.

Falcon 25-mm **cell culture inserts** from Becton Dickinson for use in six-well plates have a host of applications: growth and differentiation of polar, epithelial and

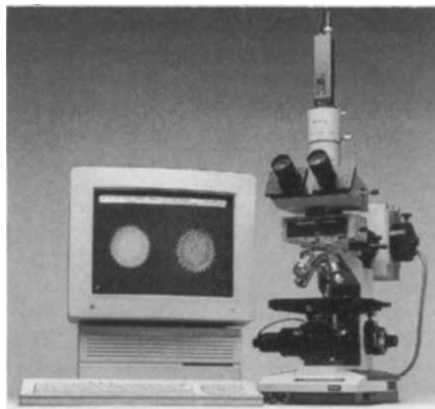


Cell culture inserts made to mimic an *in vivo* environment.

endothelial cells, *in vitro* toxicology, cell transport studies, monitoring cell-to-cell interactions, and studying viral infection, to name but a few (Reader Service No. 113). The cell culture inserts feature a special Cyclopore membrane that is designed to mimic the *in vivo* environment more closely. Pores (0.45 micron) in the membrane provide permeability to tissue culture media components, ions and low molecular weight lipoproteins. The cell culture inserts can also be used for the culture of anchorage-dependent cells. The transparent nature of the inserts makes it possible to view growing cells with phase-contrast microscopy. In addition, the membrane does not contribute to background interference when cells are treated with fluorescent stains, says the company. See booth 902.

Focusing on microscopy

Micro-Tome, an **image processing system** developed by VayTek, is designed to replace or supplement the confocal laser scanning microscope (CLSM) (Reader Service No. 114). The system consists of a plug-in array processor board and software that transforms a standard wide-field microscope, video camera and IBM AT or Apple Macintosh computer into a digital confocal microscope. A generic version of Micro-Tome is available, which works with any UNIX workstation; prices begin under \$10,000 (US). Micro-Tome



Video view with VayTek's Micro-Tome image processing system.

uses a standard white-light microscope and requires no special optical attachments. The video camera captures and digitizes the images from the microscope, which are then stored on computer. A deconvolution algorithm can be used to cut out 'noise' or 'haze' from the image. Besides increasing the resolution of the image, the deconvolved slices can be stacked to produce a three-dimensional representation of the specimen, says VayTek. Optional sectioning is non-invasive and can be performed on living specimens. For more on the Micro-Tome system, see VayTek in booth 1832.

Molecular Dynamics, who recently ac-

quired the Swedish company Sarastro, will be promoting the 1000 CLSM (**confocal laser scanning microscope**) in San Diego (*Reader Service No. 115*). The 1000 CLSM allows the user to locate and quantify specific substances within a cell or other specimens using a scanning beam of laser light to focus vertically through the cell. Powerful software and a graphics workstation can be used to reconstruct a three-dimensional representation that is displayed by the computer as a clear, sharp visual image with resolution of better than one micron, says Molecular Dynamics. Using this technology, fluorescent probes can be used to identify the precise location of genes in chromosomes. Four main elements make up the system, which sells for about \$180,000 (US): a laser scanning control module, analysis software, a standard microscope, and a Silicon Graphics workstation.

Noran Instruments will be launching Odyssey, a system that can be used to convert a standard microscope into a **laser scanning real-time confocal light microscope**, at the Cell Biology Conference (*Reader Service No. 116*). This set-up is designed primarily for use with living specimens and uses acousto-optical laser scanning to produce confocal (thin optical section) images at video rates. Odyssey has dual-channel fluorescence imaging



Noran's real-time confocal microscope.

capabilities, which offer reduced photo-bleaching and photo-toxicity effects, says Noran. Most microscopes of either standard or inverted configurations can be retrofitted with Odyssey. Two basic versions are available: the PC-version for image processing and analysis and the workstation version for three-dimensional imaging applications. See Noran in booth 613.

Carl Zeiss' range of **multi-immersion objectives** now includes new Plan-Neofluar objectives — 16×, 25× and 40× — in both brightfield and phase-contrast versions (*Reader Service No. 117*). The objectives feature a correction for oil, water or glycerine immersion, which makes them suitable for studies where 'wet' mounts are preferred and for specimens with or without coverslips, says Zeiss. The Plan-Neofluar objectives offer

transmission into the near UV. In addition, their high numerical apertures make the objectives suitable for both fluorescence microscopy and high-resolution differential contrast studies.

Catalogues/custom services

Over 500 products, including 100 new product additions, are featured in the 1991 **molecular biology catalogue** from Boehringer (*Reader Service No. 118*). An expanded section on restriction enzymes includes details on isoschizomers and 'rare' cutters. In the section on kits, users will find handy technical hints on how to streamline and improve experimental results. Boehringer's patented Genius nonradioactive nucleic acid labelling and detection technology is covered in the catalogue for the first time. See booth 421.

The Hybridoma Group at BAbCO now offers a **custom monoclonal antibody development service** (*Reader Service No. 119*). Using the company's four-phase programme — immunization, fusion/screening, cloning and antibody production — BAbCO can produce mAbs to peptide and protein molecules; viruses, bacteria and mammalian cell-surface antigens; and small organic haptens such as therapeutics, antibiotics and drugs of abuse. BAbCO also operates a custom antigen preparation service for isolating, purifying and/or conjugating an antigen. Other miscellaneous services include antibody purification, characterization, fragmentation, biotinylation, isotyping, enzyme and fluorescent labelling, and immunoassay development. BAbCO's booth is 1523 in San Diego.

Visitors to the Accurate Chemical and Scientific booth, number 1214, can pick up a copy of the updated edition of Cedarlane's **complement, monoclonal antibodies and anti-sera catalogue**



One-stop shopping with Accurate.

(*Reader Service No. 120*). The 16-page catalogue contains sections on mouse and human immunology reagents and tissue typing reagents. Information on pricing and packaging is included.

Stock up with supplies and accessories for SEM, EM and light microscopy from the 1990/1992 **catalogue** from Electron Microscopy Sciences (*Reader Service No. 121*). The new line-up features an ozone-friendly dust-off spray, an enlarging easel for making 2 × 2-inch projector slides from electronmicrographs, a two-well incubation chamber for immunocytochemistry, and new carbon adhesives. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

Serving now
BULK AMOUNTS
of
**90 RESTRICTION - and
20 MODIFYING ENZYMES**

We guarantee excellent quality
and very competitive prices.
Please inquire for quotations



AGS Angewandte Gentechnologie
Systeme GmbH
Rischerstraße 12, D-6900 Heidelberg 1
West-Germany
Tel.: Germany (6221) 81023
Fax: (6221) 840610

Reader Service No.80

DNA FINGERPRINTING

Fresenius AG, Bad Homburg is offering the following oligonucleotides **exclusively**, which were patented by Dr. J. T. Epplen, Max - Planck - Institut für Psychiatrie, Martinsried

(GTG)₅ · (GATA)₄ · (GACA)₄
(GGAT)₄ · (TCC)₅

Fresenius
Diagnostic Division

6370 Oberursel, F. R. G. • Telephone: + 61 71/60 24 43
Telefax: + 61 71/60 21 35 • Postbox 1809

Reader Service No.36

**CUSTOM-MADE & PURIFIED
DNA PEPTIDES**

\$8/base	Quotes available
50-100 µg	50-300 mg
40 hour delivery	1-3 wks delivery

GUARANTEED QUALITY & QUANTITY

V.G. CORP.

1200 Wonderland Rd. S.
Bldg #7, Unit #2
London, Ontario
Canada N6L 1A8

CALL OR FAX: (519) 652-3758

Reader Service No.29

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully.

Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50–80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in italic type) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to *Nature* are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origins and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned, most are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

GENERAL

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included. Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numerical values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained.

Figure legends must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unrefereed abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963–65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first-named.

Acknowledgements should be brief. Grant numbers and contribution numbers are not allowed.